

GenCore version 5.1.6
Copyright (C) 1993 - 2003 Compugen Ltd.

4 protein - protein search, using sw model

in on: September 29, 2003, 08:01:08 ; Search time 44 seconds
(without alignments)

1111.086 Million cell updates/sec

title: US-10-088-548-2

effect score: 1786

sequence: 1 KSPVVDYCHGDSYRGIS.....YTMNPKLFYCDIPLCASS 308

scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

sarched: 1107863 seqs, 158726573 residues

total number of hits satisfying chosen parameters: 944443

Minimum DB seq length: 0

Maximum DB seq length: 308

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A. Geneseq.19Jun03.*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1786	100.0	308	22	AA172944 Human angiogenesis
2	913	51.1	189	21	AA1918 Human plasminogen
3	913	51.1	192	21	AA1919 Human plasminogen
4	744	41.7	280	20	AA1910 A multifunctional
5	744	41.7	297	20	AA1910 A multifunctional
6	716	40.1	250	17	AA1910 Bovine kringle 1-3
7	715	40.0	250	17	AA1910 Porcine kringle 1-
8	713	39.9	268	20	AA1910 A multifunctional
9	713	39.9	268	20	AA1910 A multifunctional

10	712	39.9	260	24	AA1918 Human plasminogen
11	712	39.9	274	21	AA1918 Human plasminogen
12	711	39.8	260	21	AA1918 Deglycosylated kri
13	711	39.8	260	21	AA1918 Human Angiostatin
14	707	39.6	254	21	AA1918 Amino acid sequenc
15	704	39.4	250	17	AA1918 Rhesus kringle 1-3
16	703	39.4	250	17	AA1918 Human kringle 1-3
17	700	39.2	250	17	AA1918 Murine kringle 1-3
18	669	37.5	269	22	AA1918 Amino acid sequenc
19	585	32.8	210	20	AA1918 Anti-angiogenic an
20	539	30.2	91	22	AA1918 Human apolipoprote
21	538	30.1	168	17	AA1918 Murine kringle 2-3
22	531	29.7	168	17	AA1918 Porcine kringle 2-
23	524	29.3	130	23	AA1918 Fusion protein kri
24	523	29.3	89	22	AA1918 Human apolipoprote
25	519	29.1	91	20	AA1918 Lipoprotein. Synt
26	516	28.9	120	15	AA1918 A multifunctional
27	516	28.9	184	20	AA1918 A multifunctional
28	516	28.9	201	20	AA1918 Rhesus kringle 2-3
29	515	28.8	168	17	AA1918 Human recombinant
30	515	28.8	172	22	AA1918 Human plasminogen
31	515	28.8	175	21	AA1918 Bovine kringle 2-3
32	514	28.8	168	17	AA1918 Human kringle 2-3
33	511	28.6	168	17	AA1918 Human apolipoprote
34	508	28.4	86	22	AA1918 Bovine kringle 1-2
35	503	28.2	160	17	AA1918 Porcine kringle 1-
36	499	27.9	160	17	AA1918 Rhesus kringle 1-2
37	490	27.4	160	17	AA1918 Human kringle 1-2
38	489	27.4	160	17	AA1918 Murine kringle 1-2
39	486	27.2	160	17	AA1918 Truncated hepatocy
40	481.5	27.0	290	13	AA1918 34kd HGF variant H
41	481.5	27.0	290	14	AA1918 34 kilodalton hepa
42	481.5	27.0	290	18	AA1918 Human kringle 5 pe
43	473	26.5	101	18	AA1918 Human plasminogen
44	473	26.5	101	21	AA1918 Human plasminogen
45	473	26.5	101	21	AA1918 Human plasminogen

ALIGNMENTS

RESULT 1

AA172944

ID AA172944 standard; Protein; 308 AA.

XX AA172944;

XX 13-JUN-2001 (first entry)

XX Human angiogenesis inhibitor, LK68 protein.

XX Human; angiogenesis inhibitor; LK68; apolipoprotein(a) kringle domain;
XX angiogenesis-mediated disease; cancer; rheumatoid arthritis; therapy;
XX cytostatic; antirheumatic; antiarthritic; antipsoriatic; psoriasis;
XX ocular angiogenic disease; endothelial cell proliferation; tumour;
XX cell migration.

XX Homo sapiens.

XX Location/Qualifiers

FT Key

FT Domain

FT /label= KIV36

FT /note= "Apolipoprotein(a) kringle domain IV36, LK6

FT protein"

FT 119..207

FT /label= KIV37

FT /note= "Apolipoprotein(a) kringle domain IV37, LK7

FT protein"

FT 223..308

FT /label= KV38

FT /note= "Apolipoprotein(a) kringle domain V38, LK8

FT protein"

XX


```

Db      181 YCDVPCAA 189
      ||||| |||
RESULT 3
ID AAB01919 standard; Protein; 192 AA.
XX
AC AAB01919;
XX
DT 18-SEP-2000 (first entry)
XX
DE Human plasminogen kringle 4-5 (Val355-Phe546).
XX
KW Plasminogen; human; kringle domain; endothelial cell proliferation;
KW angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
KW antipapillary; antiinflammatory; antiulcer; antirheumatic; antiarthritic;
KW antiangiogenic; cancer; tumour; autoimmune disease.
XX
OS Homo sapiens.
XX
PN US6057122-A.
XX
PD 02-MAY-2000.
XX
PF 05-MAY-1997; 97US-0851350.
XX
PR 03-MAY-1996; 96US-0643219.
XX
PR 03-APR-1997; 97US-0832087.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Davidson DJ;
XX
DR WPI; 2000-349573/30.
XX
PT Preparation of Kringle five peptide fragment for treating various
PT disorders such as angiogenic, ocular, skin diseases and cancer,
PT involves mixing mammalian plasminogen and elastase followed by
PT incubation and isolation -
XX
PS Example 17; Page 7; 48pp; English.
XX
CC The invention relates to a method of preparing plasminogen kringle 5
CC peptide fragments. The method comprises mixing mammalian plasminogen and
CC elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC endothelial cell proliferation and migration. The peptides are useful
CC for treating angiogenic diseases, primary and metastatic solid tumours
CC and carcinomas of various organs such as breast, genital tract,
CC endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC arising from haematopoietic malignancies such as leukaemias and
CC lymphomas. They are also used for the prophylaxis of various autoimmune
CC diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC syndrome), diseases caused by excessive or abnormal stimulation of
CC endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC disease and ulcers). The peptides are also useful as a birth control
CC agent which inhibits ovulation and establishment of the placenta.
CC Sequences AAB01906-B01919 represent fragments of human plasminogen used
CC in an exemplification of the invention.
CC Note: This sequence is not shown in the specification, but is derived
CC from the full length human plasminogen sequence (AAB01887) shown in
CC figure 1.
XX
SQ Sequence 192 AA;
Query Match 51.1%; Score 913; DB 21; Length 192;
Best Local Similarity 79.4%; Pred. No. 1.3e-51;
Matches 150; Conservative 17; Mismatches 22; Indels 0; Gaps 0;
QY 119 VRQCYHGNGSQYRGFTSTVGTGTCQSSMTPHRQKTPENYPNAGLTMYNCRNPDAK 178

```

```

Db      1 VDCYHGDGQSYRGTSSTTTGKCKOSKSSMTPHRQKTPENYPNAGLTMYNCRNPDAK 60
      ||||| ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      179 GPWCFTTDPSTIRWEYCNLRCSDFIEGTVPPTVIOVPSLGPSPEDOCDFGNGKGYEGK 238
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      61 GPWCFTTDPSTIRWEYCNLRKCSGTASVAPFPVLLPDVETPSEEDOCDFGNGKGYEGK 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      239 ATTGTGTCQDWAQEPHRSFTTFTGINKWAGLEKYNCRNPDGIDNGPWCYTNNPKLFD 298
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      121 ATTGTGTCQDWAQEPHRSFTTFTGINKWAGLEKYNCRNPDGIDNGPWCYTNNPKLFD 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      299 YCDIPLCAS 307
      ||||| |||
Db      181 YCDVPCAA 189
      ||||| |||
RESULT 4
ID AAV02108 standard; Protein; 280 AA.
XX
AC AAV02108;
XX
DT 16-JUL-1999 (first entry)
XX
DE A multifunctional protein of the invention.
XX
KW Angiostatin; endostatin; interferon; thrombospondin;
KW interferon-inducible protein; platelet factor 4; anti-angiogenic;
KW anti-tumor; multifunctional protein; angiogenic-mediated disease;
KW cancer; diabetic retinopathy; macular degeneration; arthritis;
KW tumor cell production.
XX
OS Synthetic.
XX
OS Homo sapiens.
XX
PN WO9916889-A1.
XX
PD 08-APR-1999.
XX
PF 30-SEP-1998; 98WO-US20464.
XX
PR 01-OCT-1997; 97US-0060609.
XX
PA (SEAR ) SEARLE & CO G D.
XX
PI Bolanowski MA, Caparon MH, Casperson GF, Gregory SA;
PI Klein BK, McKearn JP;
XX
DR WPI; 1999-255098/21.
XX
PT New multifunctional proteins useful for treating angiogenic-mediated
PT diseases
XX
CC Claim 5; Page 103-104; 121pp; English.
XX
CC The specification describes multifunctional proteins which comprise
CC combinations of angiostatin, endostatin, interferon, thrombospondin,
CC interferon-inducible protein and platelet factor 4, and have
CC anti-angiogenic and/or anti-tumor activity. The multifunctional protein
CC may exhibit useful properties such as having similar or greater
CC biological activity when compared to a single factor or by having
CC improved half-life or decreased adverse side effects, or a combination
CC of these properties. The proteins can be used for treating an
CC angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular
CC degeneration, or arthritis. They can also be used for inhibiting the
CC production of tumor cells (characteristic of lung, breast, ovarian,
CC prostate, pancreatic, gastric, colon, renal, bladder cancers; melanoma,
CC hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor
CC growth. The present sequence represents a multifunctional protein of the
CC invention.
XX
SQ Sequence 280 AA;

```


PT Use of angiostatin fragments or aggregates - for inhibiting
 PT endothelial cell proliferation and treating angiogenesis-mediated
 PT diseases, e.g. cancer, arthritis or diabetic retinopathy
 XX
 XX
 PS Claim 4; Page 128; 203pp; English.
 XX
 CC The invention relates to new methods and compositions for
 CC inhibiting endothelial cell proliferation, using as active component
 CC an angiostatin fragment, a combination of angiostatin fragments, or
 CC aggregate angiostatin. The fragment is preferably derived from murine,
 CC human, rhesus, porcine or bovine plasminogen and is a kringle 1,
 CC kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle
 CC 1-4 or kringle 1-4BKLs protein. The aggregate angiostatin has a Mol. Wt.
 CC of 45-65 kD and is derived from a plasminogen fragment beginning at
 CC approximately amino acid number 98 of murine, human, Rhesus, porcine or
 CC bovine plasminogen. The active component can be used for treating
 CC angiogenesis-mediated diseases such as cancer, arthritis, macular
 CC degeneration and diabetic retinopathy. It can also be used to develop
 CC antibodies for use in diagnosis, detection and therapy.
 CC The present sequence, bovine kringle 1-3, is a specific angiostatin
 CC fragment which can be used in the invention, and represents amino
 CC acids 6-255 of bovine angiostatin.
 XX
 SQ Sequence 250 AA;

Query Match 40.1%; Score 716; DB 17; Length 250;
 Best Local Similarity 42.1%; Pred. No. 8.6e-39;
 Matches 126; Conservative 43; Mismatches 80; Indels 50; Gaps 7;

QY 8 CYHGDGRYRGISSTVTGTCQSWSSMIPHWORTPENYPNAGLTENYCRNPDSGQ-P 66
 Db 1 CTGNGQYRGTTAETKSGVTCQKNSATSPVVPKFSPEKPLAGLENYCRNPNDENG 60

QY 67 WCYTDDPCVWEYCNLTQCSSETSGVLETPVVPVPSMEAHSEAPTEQTPVVRQCYHGN 126
 Db 61 WCYTDDPCVWEYCNLTQCSSETSGVLETPVVPVPSMEAHSEAPTEQTPVVRQCYHGN 126

QY 127 GQSYRGTSFTVTGTCQSWSSMTPHQRTPENYPNDGLTMVYCRNPDPADTGPWCFTTD 186
 Db 88 GENYEGKIATKMSGDCQANDSQSPHAGVYPSKFPNKNLKNYCRNPDPGPRPWCFTTD 147

QY 187 PSIRWEYCNLTQCSSETSGVLETPVVPVPSMEAHSEAPTEQTPVVRQCYHGN 246
 Db 148 POKRWEFCIDIPRCT-----TPP-----PSSGP--KYQCLKGTGKNYGGTVAVTESGHT 193

QY 247 QEWAAQEPHRSHTFIPGTNKGLEKNYCRNPDPGDPINGPWCYTMNPKLFDCYDIPLC 305
 Db 194 QCRWSEQPQHNR-TPENFPCKNLEENYCRNPDPGE-KAPWCITTNSEVRWEICITPSC 250

RESULT 7
 AA07569
 ID AA07569 standard; protein; 250 AA.

XX
 AC AA07569;
 XX
 XX 22-JUN-1997 (first entry)

XX Porcine kringle 1-3.

XX angiostatin; plasminogen; kringle; angiogenesis; cancer; arthritis;
 XX macular degeneration; diabetic retinopathy.

XX Sus scrofa.

XX WO9635774-A2.

XX 14-NOV-1996.

XX 26-APR-1996; 96WO-US05856.

XX 08-MAR-1996; 96US-0612788.

PR 26-APR-1995; 95US-0429743.

PR 22-FEB-1996; 96US-0605598.
 XX (CHIL-) CHILDRENS MEDICAL CENT.
 XX
 PI Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;
 XX WPI; 1996-518662/51.
 DR
 XX
 XX Use of angiostatin fragments or aggregates - for inhibiting
 PT endothelial cell proliferation and treating angiogenesis-mediated
 PT diseases, e.g. cancer, arthritis or diabetic retinopathy
 XX
 PS Claim 4; Page 126-127; 203pp; English.

XX The invention relates to new methods and compositions for
 CC inhibiting endothelial cell proliferation, using as active component
 CC an angiostatin fragment, a combination of angiostatin fragments, or
 CC aggregate angiostatin. The fragment is preferably derived from murine,
 CC human, Rhesus, porcine or bovine plasminogen and is a kringle 1,
 CC kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle
 CC 1-4 or kringle 1-4BKLs protein. The aggregate angiostatin has a Mol. Wt.
 CC of 45-65 kD and is derived from a plasminogen fragment beginning at
 CC approximately amino acid number 98 of murine, human, Rhesus, porcine or
 CC bovine plasminogen. The active component can be used for treating
 CC angiogenesis-mediated diseases such as cancer, arthritis, macular
 CC degeneration and diabetic retinopathy. It can also be used to develop
 CC antibodies for use in diagnosis, detection and therapy.
 CC The present sequence, porcine kringle 1-3, is a specific angiostatin
 CC fragment which can be used in the invention, and represents amino
 CC acids 6-255 of porcine angiostatin.
 XX
 SQ Sequence 250 AA;

Query Match 40.0%; Score 715; DB 17; Length 250;
 Best Local Similarity 42.5%; Pred. No. 1e-38;
 Matches 127; Conservative 41; Mismatches 81; Indels 50; Gaps 7;

QY 8 CYHGDGRYRGISSTVTGTCQSWSSMIPHWORTPENYPNAGLTENYCRNPDSG-KOP 66
 Db 1 CTGNGKNYRGTSKTSKSVTCQKNSVSSPHIKYSPFAFPLAGLENYCRNPNDKEGP 60

QY 67 WCYTDDPCVWEYCNLTQCSSETSGVLETPVVPVPSMEAHSEAPTEQTPVVRQCYHGN 126
 Db 61 WCYTDDPCVWEYCNLTQCSSETSGVLETPVVPVPSMEAHSEAPTEQTPVVRQCYHGN 126

QY 127 GQSYRGTSFTVTGTCQSWSSMTPHQRTPENYPNDGLTMVYCRNPDPADTGPWCFTTD 186
 Db 88 GEHTEGKISMTSGIECQSGSQSPHAGVYPSKFPNKNLKNYCRNPDPGPRPWCFTTD 147

QY 187 PSIRWEYCNLTQCSSETSGVLETPVVPVPSMEAHSEAPTEQTPVVRQCYHGN 246
 Db 148 POKRWEFCIDIPRCT-----TPP-----PSSGP--KYQCLKGTGKNYGGTVAVTESGHT 193

QY 247 QEWAAQEPHRSHTFIPGTNKGLEKNYCRNPDPGDPINGPWCYTMNPKLFDCYDIPLC 305
 Db 194 QCRWSEQPQHNR-TPENFPCKNLEENYCRNPDPGE-KAPWCITTNSEVRWEICITPSC 250

RESULT 8

AA02109
 ID AA02109 standard; Protein; 268 AA.

XX
 AC AA02109;

XX
 DT 16-JUL-1999 (first entry)

XX A multifunctional protein of the invention.

XX Angiostatin; endostatin; interferon; thrombospondin;
 KW interferon-inducible protein; platelet factor 4; anti-angiogenic;
 KW anti-tumor; multifunctional protein; angiogenic-mediated disease;
 KW cancer; diabetic retinopathy; macular degeneration; arthritis;
 KW tumor cell production.

1 Synthetic.
2 Homo sapiens.
3 WO9916889-A1.
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

AA02102 standard; Protein; 285 AA.
AA02102;
16-JUL-1999 (first entry)
A multifunctional protein of the invention.
Angiostatin; endostatin; interferon; thrombospondin;
interferon-inducible protein; platelet factor 4; anti-angiogenic;
anti-tumor; multifunctional protein; angiogenic-mediated disease;
cancer; diabetic retinopathy; macular degeneration; arthritis;
tumor cell production.
Synthetic.
Homo sapiens.
WO9916889-A1.
08-APR-1999.
30-SEP-1998; 98WO-US20464.
01-OCT-1997; 97US-0060609.
(SEAR) SEARLE & CO G D.
Bolnowski MA, Caparon MH, Casperson GF, Gregory SA;
Klein BK, McKearn JP;
WPI; 1999-255098/21.
New multifunctional proteins useful for treating angiogenic-mediated
diseases
Claim 5; Page 97-98; 121pp; English.
The specification describes multifunctional proteins which comprise
combinations of angiostatin, endostatin, interferon, thrombospondin,
interferon-inducible protein and platelet factor 4, and have
may exhibit useful properties such as having similar or greater
biological activity when compared to a single factor or by having
improved half-life or decreased adverse side effects, or a combination
of these properties. The proteins can be used for treating an
angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular
degeneration, or arthritis. They can also be used for inhibiting the
production of tumor cells (characteristic of lung, breast, ovarian,
prostate, pancreatic, gastric, colon, renal, bladder cancers; melanoma,
hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor
growth. The present sequence represents a multifunctional protein of the
invention.
Sequence 285 AA;
Query Match 39.9%; Score 713; DB 20; Length 285;
Best Local Similarity 42.2%; Pred. No. 1.4e-38;
Matches 130; Conservative 41; Mismatches 87; Indels 50; Gaps 7;
2 SPVQDCYHGDGRGSRGTSSTVYGTGTCQSWSMTPHWHQRTPENYNAGLTENYCRNPD 61
2 SVYLSECKTGNGKNYRGTMTSKNGITCQKWSSTSPHRFPSPATHPSEGLENYCRNPD 61
62 SGKQ-PWCYTTPDCVRWEYCNLTQCSSTESGVLETPTVVPVPSMEAHSEAAPEQTTPVVR 120
62 NDFQGFWCITTPDKRYDCILECE----- 88
121 QCYHNGQSYRGTFSTTVGTGTCQSWSMTPHWHQRTPENYNAGLTENYCRNPDADTGP 180
89 ECMHCSGSEYDCKISKTSMSGLEQAWDSQSPHAGYIPSKFPNKLKKNYCRNPDRLEP 148
181 WCFTTDPSTRWEYCNLTQCSSTEGTVVAPPTVQVPSLGPSPSEODCMFGNGKYRGKAT 240
149 WCFTTDPNKRWLCDIPRCT-----TPP-----PSSGPTYQ--CLKGTGNYRGNVAV 194
241 TVTGTTCQEWAAQEPHRTSTIPGNTKWAGLEKNYCRNPDPDINGPWCYTNPDKLFYDC 300
195 TVSHGTCQHWSAQTPTHTNR--TPNFCKNLNENYCRNPDKG-RAPWCHTNTNSQVRWEYC 252
301 DIPLCASS 308
253 KIPSCDSS 260
RESULT 9
AA02102

X T Purifying recombinant angiostatin, involves applying fermented broth
 T containing angiostatin to expanded bed cation exchange column, anion
 T exchange column, hydroxyapatite column, hydrophobic column and a
 T membrane
 X S
 X Example 1; Page 22; 49pp; English.
 X C The present sequence representing Angiostatin protein is given in an
 C invention providing a method for recombinant production, recovery and
 C purification of Angiostatin protein. Purification of recombinant
 C Angiostatin comprises applying crude fermentation broth containing the
 C protein to an expanded bed cation exchange column, eluting it, and
 C applying the eluate to anion exchange column, repeating the process of
 C eluting and applying, to hydroxyapatite column, hydrophobic column and
 C membrane, in order, and collecting fluid passing through the membrane.
 C Angiostatin is useful for treating angiogenesis mediated diseases,
 C including solid tumours, leukaemia, tumour metastases, benign tumours,
 C rheumatoid arthritis, psoriasis, ocular angiogenic diseases, Osler-Webber
 C syndrome, myocardial angiogenesis, plaque neovascularisation,
 C telangiectasia, haemophilic joints, angiofibroma and wound granulation.
 C As a centrifugation technique is not employed in the process, damage to
 C the cells with concomitant release of undesirable biological materials
 C such as cytochromes, pigments, enzymes, chemicals and other undesirable
 C cellular constituents and debris, is prevented. Large scale recovery and
 C purification of proteins is greater than that obtained from prior art
 C methods. Active Angiostatin can be stored in buffers for extended periods
 C of time, in vials or other containers, either in solution which may be
 C liquid or frozen, or lyophilised.
 X Q Sequence 260 AA;

Query Match 39.8%; Score 711; DB 22; Length 260;
 Best Local Similarity 42.3%; Pred. No. 1.9e-38;
 Matches 129; Conservative 40; Mismatches 86; Indels 50; Gaps 7;
 Y 5 VQCYHGDGRSIRGTSITVGRFCOSWSMIPHWQRTPEPNAGLTENYCRNPDGK 64
 b 1 LSECKTGNKNGYGTMSKTKNGITCQKWSSTSPHPRSPATHSEGLEENYCRNPDNDP 62
 Y 65 Q-PWCYTTPDCVRYWYCNLQCSSTESGVLETPVVPVPSMEAHSEAPTEQTPVRCY 123
 b 63 QGPWCYTTPDCVRYWYCNLQCSSTESGVLETPVVPVPSMEAHSEAPTEQTPVRCY 89
 Y 124 HGNGQSYRGTFSTVGRFCOSWSMIPHWQRTPEPNAGLTENYCRNPDADGPKCF 183
 b 90 HCSGNYDGRISKTSGLEQAWDSQSPHAGYIPSKFPNKLKKNYCRNPDRELPRWCF 149
 Y 184 TTPDSIRWEYCNLTRCSDTEGTVVAPPTVIOVPSLGPSPSEQDMFGNGKGRKATVT 243
 b 150 TTPDNKRWELCDIPRCT-----TPP-----PSSGPTYQ--CLKGTGENYRGNVAVTS 195
 Y 244 GTCQWAAQEPHRSHTFIPGTNKNWAGLEKNYCRNPDGNGPCWCTMNPRLKFDYCDIP 303
 b 196 GHTCQWAAQEPHRSHTFIPGTNKNWAGLEKNYCRNPDGNGPCWCTMNPRLKFDYCDIP 253
 Y 304 LCASS 308
 b 254 SCDS 258
 RESULT 14
 AY33869
 D AAY53869 standard; protein; 254 AA.
 C AAY53869;
 X Amino acid sequence of human greenstatin protein.
 T 13-MAR-2000 (first entry)
 X Human; plasminogen; angiostatin; greenstatin; thrombolytic factor;
 W angiogenesis inhibitory protein; proliferation; angiogenesis; cancer;
 X

KW vascular endothelial cell; ophthalmic disease; glaucoma;
 KW diabetic retinopathy; arthritis; psoriasis.
 XX Homo sapiens.
 OS WO9961464-A1.
 PN 02-DEC-1999.
 PD 28-MAY-1999; 99WO-KR00263.
 PP 28-MAY-1998; 98KR-0019335.
 PR 27-MAY-1999; 99KR-0019144.
 PA (GREC) KOREA GREEN CROSS CORP.
 PI You WK, So SH, Ahn BC, Lee H, Jung S, Kim Y, Lee JH, Hong Y;
 PI Joe YA, Chang S;
 XX WPI; 2000-086703/07.
 XX Purifying angiogenesis inhibitors produced as recombinant proteins in
 PT Escherichia coli, useful as anticancer agents and for treating ocular
 PT diseases
 XX Claim 2; Page 47-48; 55pp; English.
 CC The present sequence represents the human greenstatin protein. It
 CC is derived from the plasminogen protien, and comprises amino acids
 CC 101-354. Angiostatin is also derived from plasminogen, and comprises
 CC amino acids 99-467. Angiostatin and greenstatin are used as
 CC thrombolytic factors and angiogenesis inhibitory proteins. Angiostatin
 CC contains the kringle 1-4 region of plasminogen, and greenstatin contains
 CC the kringle 1-3 region of plasminogen. As both proteins contain a high
 CC number of disulphide bonds, they are difficult to purify. The
 CC specification describes a method for the purification of such
 CC angiogenesis inhibitory proteins. The method comprises solubilising
 CC the proteins, produced as inclusion bodies in Escherichia coli and
 CC refolding the solubilised fraction in buffer containing urea and
 CC glutathione. The angiogenesis inhibitory proteins specifically inhibit
 CC proliferation of vascular endothelial cells, but not that of
 CC non-endothelial cancers or normal cells. The angiogenesis inhibitory
 CC proteins are used to suppress angiogenesis, specifically for treating
 CC cancers (e.g. of lung, skin and brain) or ophthalmic diseases (e.g.
 CC glaucoma and diabetic retinopathy), but also arthritis and psoriasis.
 XX Sequence 254 AA;
 SQ
 Query Match 39.6%; Score 707; DB 21; Length 254;
 Best Local Similarity 42.4%; Pred. No. 3.3e-38;
 Matches 128; Conservative 40; Mismatches 84; Indels 50; Gaps 7;
 QY 7 DCYHGDGRSIRGTSITVGRFCOSWSMIPHWQRTPEPNAGLTENYCRNPDGSKQ- 65
 DB 2 ECKTGNKNGYGTMSKTKNGITCQKWSSTSPHPRSPATHSEGLEENYCRNPDNDPQ 61
 QY 66 PWCYTTPDCVRYWYCNLQCSSTESGVLETPVVPVPSMEAHSEAPTEQTPVRCYCHG 125
 DB 62 PWCYTTPDCVRYWYCNLQCSSTESGVLETPVVPVPSMEAHSEAPTEQTPVRCYCHG 88
 QY 126 NGQSYRGTFSTVGRFCOSWSMIPHWQRTPEPNAGLTENYCRNPDADGPKCF 185
 DB 89 SGENYDGRISKTSGLEQAWDSQSPHAGYIPSKFPNKLKKNYCRNPDRELPRWCF 148
 QY 186 DPSIRWEYCNLTRCSDTEGTVVAPPTVIOVPSLGPSPSEQDMFGNGKGRKATVTGT 245
 DB 149 DPNKRWELCDIPRCT-----TPP-----FSSGPTYQ--CLKGTGENYRGNVAVTS 194
 QY 246 PCQWAAQEPHRSHTFIPGTNKNWAGLEKNYCRNPDGNGPCWCTMNPRLKFDYCDIP 305
 DB 195 TQCHWSAQTPTHTNR-TPEFNPKNLNENYCRNPDGK-RAPWCHTNSQVWREYKIPSC 252
 QY 306 AS 307

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:01:08 ; Search time 29 seconds
(without alignments)
449.370 Million cell updates/sec

Title: US-10-088-548-2

Perfect score: 1786

Sequence: 1 KSPVVDYHGDGRSYRIS.....YTMNPKLFYCDIPLCASS 308

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 282628

Minimum DB seq length: 0

Maximum DB seq length: 308

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCFUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	716	40.1	250	2	US-08-612-788-33
2	716	40.1	250	3	US-09-066-028-33
3	716	40.1	250	4	US-09-335-325-33
4	715	40.0	250	2	US-08-612-788-32
5	715	40.0	250	3	US-09-066-028-32
6	715	40.0	250	4	US-09-335-325-32
7	704	39.4	250	2	US-08-612-788-31
8	704	39.4	250	3	US-09-066-028-31
9	704	39.4	250	4	US-09-335-325-31
10	703	39.4	250	2	US-08-612-788-30
11	703	39.4	250	3	US-09-066-028-30
12	703	39.4	250	4	US-09-335-325-30
13	700	39.2	250	2	US-08-612-788-29
14	700	39.2	250	3	US-09-066-028-29
15	700	39.2	250	4	US-09-335-325-29
16	585	32.8	210	3	US-08-985-526-21
17	538	30.1	168	2	US-08-612-788-24
18	538	30.1	168	3	US-09-066-028-24
19	538	30.1	168	4	US-09-335-325-24
20	531	29.7	168	2	US-08-612-788-27
21	531	29.7	168	3	US-09-066-028-27
22	531	29.7	168	4	US-09-335-325-27
23	524	29.3	130	4	US-08-240-839-4
24	519	29.1	91	3	US-09-234-553-1
25	516	28.9	120	1	US-08-211-747-7
26	515	28.8	168	2	US-08-612-788-26
27	515	28.8	168	3	US-09-066-028-26

28	515	28.8	168	4	US-09-335-325-26
29	514	28.8	168	2	US-08-612-788-28
30	514	28.8	168	3	US-09-066-028-28
31	514	28.8	168	4	US-09-335-325-28
32	511	28.6	168	2	US-08-612-788-25
33	511	28.6	168	3	US-09-066-028-25
34	511	28.6	168	4	US-09-335-325-25
35	503	28.2	160	3	US-08-612-788-38
36	503	28.2	160	4	US-09-066-028-38
37	503	28.2	160	2	US-09-335-325-38
38	499	27.9	160	3	US-08-612-788-37
39	499	27.9	160	4	US-09-066-028-37
40	499	27.9	160	2	US-09-335-325-37
41	490	27.4	160	4	US-08-612-788-36
42	490	27.4	160	3	US-09-066-028-36
43	490	27.4	160	2	US-09-335-325-36
44	489	27.4	160	2	US-08-612-788-35
45	489	27.4	160	3	US-09-066-028-35

ALIGNMENTS

RESULT 1
US-08-612-788-33
; Sequence 33, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Bovine
; IMMEDIATE SOURCE:
; CLONE: KI-3
; US-08-612-788-33

Sequence 26, Appl
Sequence 28, Appl
Sequence 28, Appl
Sequence 28, Appl
Sequence 25, Appl
Sequence 25, Appl
Sequence 38, Appl
Sequence 38, Appl
Sequence 37, Appl
Sequence 37, Appl
Sequence 36, Appl
Sequence 36, Appl
Sequence 35, Appl
Sequence 35, Appl


```

/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE:
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 31:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 250 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Rhesus monkey
/ IMMEDIATE SOURCE:
/ CLONE: KI-3
/ US-08-612-788-31

Query Match 39.4%; Score 704; DB 2; Length 250;
Best Local Similarity 42.5%; Pred. No. 2.2e-56;
Matches 127; Conservative 40; Mismatches 82; Indels 50; Gaps 7;

QY 8 CYHGDGRSGYRGSSITVTRGTCQSWSMIPHWORTPENYPNAGLTENYCRNPDG-KGOP 66
Db 1 CKTGNGKNYRGVMSKTRGTCQKWSSTSPHRFTFSPATHPSEGLEENYCRNPDNDGQGP 60
QY 67 WCYTTPDCVRWEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAPTEQTPVVRQCYHGN 126
Db 61 WCYTTPDEERFYCDIPECED-----ECMHCS 87
QY 127 GQSYRGFTFTVTRGTCQSWSMTPHRHQRTPENYPNDGLTMYNCRNPDADTGPWCFTTD 186
Db 88 GENDGKISKTSMGLEQAWDSQSPHAGHYIPSKFPNKLKNYCRNPDGEPRPWCFTTD 147
QY 187 PSIRWEYCNLTGCSDEGTGVVAPPTVIQVPSLGPSPSEQDCMGNGKGYRGKATTVTGP 246
Db 148 PNKWELCDIPRCT-----TPP-----PSSGPTYQ--CLAGTGENYRGDVAVTVSGHT 193
QY 247 CQEWAAQEPHRHSTPIPTGNKAGLEKNYCRNPDGDNFGWCIYMNPKLFDYCDIPLC 305
Db 194 CHGWSAQTPHTNR-TPENFFCKNLNENYCRNPDGE-KAPWCYTTNSQVRWEYCKIPSC 250

RESULT 8
US-09-066-028-31
/ Sequence 31, Application US/09066028
/ Patent No. 6024688
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ APPLICANT: O'Reilly, Micheal
/ APPLICANT: Cao, Yihai
/ APPLICANT: Sim, B. Kim Lee
/ TITLE OF INVENTION: Angiostatin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S.
/ ZIP: 30303-1769
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/066,028
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/612,788
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 31:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 250 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Rhesus monkey
/ IMMEDIATE SOURCE:
/ CLONE: KI-3
/ US-09-066-028-31

Query Match 39.4%; Score 704; DB 3; Length 250;
Best Local Similarity 42.5%; Pred. No. 2.2e-56;
Matches 127; Conservative 40; Mismatches 82; Indels 50; Gaps 7;

QY 8 CYHGDGRSGYRGSSITVTRGTCQSWSMIPHWORTPENYPNAGLTENYCRNPDG-KGOP 66
Db 1 CKTGNGKNYRGVMSKTRGTCQKWSSTSPHRFTFSPATHPSEGLEENYCRNPDNDGQGP 60
QY 67 WCYTTPDCVRWEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAPTEQTPVVRQCYHGN 126
Db 61 WCYTTPDEERFYCDIPECED-----ECMHCS 87
QY 127 GQSYRGFTFTVTRGTCQSWSMTPHRHQRTPENYPNDGLTMYNCRNPDADTGPWCFTTD 186
Db 88 GENDGKISKTSMGLEQAWDSQSPHAGHYIPSKFPNKLKNYCRNPDGEPRPWCFTTD 147
QY 187 PSIRWEYCNLTGCSDEGTGVVAPPTVIQVPSLGPSPSEQDCMGNGKGYRGKATTVTGP 246
Db 148 PNKWELCDIPRCT-----TPP-----PSSGPTYQ--CLAGTGENYRGDVAVTVSGHT 193
QY 247 CQEWAAQEPHRHSTPIPTGNKAGLEKNYCRNPDGDNFGWCIYMNPKLFDYCDIPLC 305
Db 194 CHGWSAQTPHTNR-TPENFFCKNLNENYCRNPDGE-KAPWCYTTNSQVRWEYCKIPSC 250

RESULT 9
US-09-335-325-31
/ Sequence 31, Application US/09335325
/ Patent No. 6521439
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ APPLICANT: O'Reilly, Micheal
/ APPLICANT: Cao, Yihai
/ APPLICANT: Sim, B. Kim Lee
/ TITLE OF INVENTION: Angiostatin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S.

```

```

;
; ZIP: 30303-1769
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
;
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Rhesus monkey
; IMMEDIATE SOURCE:
; CLONE: K1-3
; SEQUENCE DESCRIPTION: SEQ ID NO: 31:

```

```

US-09-335-325-31

```

```

Query Match 39.4%; Score 704; DB 4; Length 250;
Best Local Similarity 42.5%; Pred. No. 2.2e-56;
Matches 127; Conservative 40; Mismatches 82; Indels 50; Gaps 7;

QY 8 CYHGRSIRGSSITVTGTCQSWSMTPHQRHPTPENYPNAGLTENYCRNPDGSGKQ-P 66
Db 1 CKTGKNGKRGMTSRTGTCQKWSSTSPHRTSPATHPSEGLENYCRNPDNDGQGP 60
QY 67 WCYTDPQVRWEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAAPEQTPTVVCQYHGN 126
Db 61 WCYTDPQVRWEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAAPEQTPTVVCQYHGN 126
QY 127 GQSYRGTFSTVTGTCQSWSMTPHQRHPTPENYPNAGLTENYCRNPDADTGPWCFTTD 186
Db 88 GENTYDGIKSTMSGLEQANDSQSPAHGYIPSKFPNKLKKNYCRNPDRELPRWCFTTD 147
QY 187 PSIRWEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAAPEQTPTVVCQYHGN 246
Db 148 PNKRWELCDIPRCT-TPP-----PSSGPTYQ-CLKGTGENYRGVAVTVSGHT 193
QY 247 CQWAAQEPHRSSTIPGTPNKGAGLENYCRNPDGDPGMPWCTMNPRLFDYCDIPLC 305
Db 194 CHGWSAQTPHNR-TPENFPCKNLNENYCRNPDGE-KAPWCYTTSQVWEYCKIPSC 250

RESULT 10
US-08-612-788-30
; Sequence 30, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45

```

```

;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: K1-3
;
; US-08-612-788-30

```

```

Query Match 39.4%; Score 703; DB 2; Length 250;
Best Local Similarity 42.5%; Pred. No. 2.7e-56;
Matches 127; Conservative 39; Mismatches 83; Indels 50; Gaps 7;

QY 8 CYHGRSIRGSSITVTGTCQSWSMTPHQRHPTPENYPNAGLTENYCRNPDGSGKQ-P 66
Db 1 CKTGKNGKRGMTSRTGTCQKWSSTSPHRTSPATHPSEGLENYCRNPDNDGQGP 60
QY 67 WCYTDPQVRWEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAAPEQTPTVVCQYHGN 126
Db 61 WCYTDPQVRWEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAAPEQTPTVVCQYHGN 126
QY 127 GQSYRGTFSTVTGTCQSWSMTPHQRHPTPENYPNAGLTENYCRNPDADTGPWCFTTD 186
Db 88 GENTYDGIKSTMSGLEQANDSQSPAHGYIPSKFPNKLKKNYCRNPDRELPRWCFTTD 147
QY 187 PSIRWEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAAPEQTPTVVCQYHGN 246
Db 148 PNKRWELCDIPRCT-TPP-----PSSGPTYQ-CLKGTGENYRGVAVTVSGHT 193
QY 247 CQWAAQEPHRSSTIPGTPNKGAGLENYCRNPDGDPGMPWCTMNPRLFDYCDIPLC 305
Db 194 CHGWSAQTPHNR-TPENFPCKNLNENYCRNPDGE-KAPWCYTTSQVWEYCKIPSC 250

RESULT 11
US-09-066-028-30
; Sequence 30, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee

```



```

RESULT 13
US-08-612-788-29
; Sequence 29, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K1-3
US-08-612-788-29

Query Match 39.28; Score 700; DB 2; Length 250;
Best Local Similarity 41.18; Pred. No. 5e-56;
Matches 123; Conservative 37; Mismatches 89; Indels 50; Gaps 6;

QY 8 CYHGGRSYRGISSTVTGRTQCSWSSMIPRHQRTPEYVNPAGLTENYCRNPDGSKQ-P 66
Db 1 CKTGIGYRGYMTSKTSGVACQKAGATFPHVPVNSPTHTPEGLNENYCRNPDNDGQGP 60
QY 67 WCTTDPVCRWEYCNLQCSSETESGVLETPVTPVPSMEASAPTEQTPVVRQCYHGN 126
Db 61 WCYTDPDKRYDNCNPECEE-----ECMYCS 87
QY 127 GQSYRGTFTVTGRTQCSWSSMIPRHQRTPEYVNPAGLTENYCRNPDADTGWCTTTD 186
Db 88 GEYEGKISKTMISGLDQANDSOPHAGYIPAKFPKSKNLMKYNCHNPDGPRWCFTTD 147
QY 187 PSIRWEYCNLRCSDEGTGVVAPTVIQVPSLGPFSQDCMFGNGKGYRGKATTVTGT 246
Db 148 PTKRWYECIDPRC-----TTPPPP-----PSPTYQLKGRGNRGVTSVTSK 193
QY 247 CQWRAQPHRHHSIFGTHKAGLEKYNCRNPDGDNPGWCTMTNPKLFYCDIFLC 305

```

```

Db 194 CQWSEQTPHNR-TPENPCKNLEENYCRNPDGE-TAPWCYTDDSQIRWEYCEIPSC 250

RESULT 14
US-09-066-028-29
; Sequence 29, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K1-3
US-09-066-028-29

Query Match 39.28; Score 700; DB 3; Length 250;
Best Local Similarity 41.18; Pred. No. 5e-56;
Matches 123; Conservative 37; Mismatches 89; Indels 50; Gaps 6;

QY 8 CYHGGRSYRGISSTVTGRTQCSWSSMIPRHQRTPEYVNPAGLTENYCRNPDGSKQ-P 66
Db 1 CKTGIGYRGYMTSKTSGVACQKAGATFPHVPVNSPTHTPEGLNENYCRNPDNDGQGP 60
QY 67 WCTTDPVCRWEYCNLQCSSETESGVLETPVTPVPSMEASAPTEQTPVVRQCYHGN 126
Db 61 WCYTDPDKRYDNCNPECEE-----ECMYCS 87
QY 127 GQSYRGTFTVTGRTQCSWSSMIPRHQRTPEYVNPAGLTENYCRNPDADTGWCTTTD 186
Db 88 GEYEGKISKTMISGLDQANDSOPHAGYIPAKFPKSKNLMKYNCHNPDGPRWCFTTD 147

```


GenCore version 5.1.6

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:03:07 : Search time 347 Seconds

(without alignments)
134.301 Million cell updates/sec

Title: US-10-088-548-2

Perfect score: 1786

Sequence: 1 KSPVVQDCYHGGRSYRGIS.....YTMNPKLFYCDIPLCASS 308

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 373251

Minimum DB seq length: 0

Maximum DB seq length: 308

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2.6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2.6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2.6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2.6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2.6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2.6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2.6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2.6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2.6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2.6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2.6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2.6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2.6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2.6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2.6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2.6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2.6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2.6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	716	40.1	250	9	US-09-761-120-33
2	716	40.1	250	10	US-09-335-325-33
3	716	40.1	250	15	US-10-131-241-33
4	715	40.0	250	9	US-09-761-120-32
5	715	40.0	250	10	US-09-335-325-32
6	715	40.0	250	15	US-10-131-241-32
7	711	39.8	250	15	US-10-131-241-61
8	704	39.4	250	9	US-09-761-120-31
9	704	39.4	250	10	US-09-335-325-31
10	704	39.4	250	15	US-10-131-241-31
11	703	39.4	250	9	US-09-761-120-30
12	703	39.4	250	10	US-09-335-325-30
13	703	39.4	250	15	US-10-131-241-30
14	700	39.2	250	9	US-09-761-120-29
15	700	39.2	250	10	US-09-335-325-29

16	700	39.2	250	15	US-10-131-241-29	Sequence 29, Appl
17	585	32.8	210	14	US-10-036-889-21	Sequence 21, Appl
18	538	30.1	168	9	US-09-761-120-24	Sequence 24, Appl
19	538	30.1	168	10	US-09-335-325-24	Sequence 24, Appl
20	538	30.1	168	15	US-10-131-241-24	Sequence 24, Appl
21	531	29.7	168	9	US-09-761-120-27	Sequence 27, Appl
22	531	29.7	168	10	US-09-335-325-27	Sequence 27, Appl
23	531	29.7	168	15	US-10-131-241-27	Sequence 27, Appl
24	519	29.1	91	9	US-09-748-468-1	Sequence 1, Appl
25	515	28.8	168	9	US-09-761-120-26	Sequence 26, Appl
26	515	28.8	168	10	US-09-335-325-26	Sequence 26, Appl
27	515	28.8	168	15	US-10-131-241-26	Sequence 26, Appl
28	514	28.8	168	9	US-09-761-120-28	Sequence 28, Appl
29	514	28.8	168	10	US-09-335-325-28	Sequence 28, Appl
30	514	28.8	168	15	US-10-131-241-28	Sequence 28, Appl
31	511	28.6	168	9	US-09-761-120-25	Sequence 25, Appl
32	511	28.6	168	10	US-09-335-325-25	Sequence 25, Appl
33	511	28.6	168	15	US-10-131-241-25	Sequence 25, Appl
34	503	28.2	160	9	US-09-761-120-38	Sequence 38, Appl
35	503	28.2	160	10	US-09-335-325-38	Sequence 38, Appl
36	503	28.2	160	15	US-10-131-241-38	Sequence 38, Appl
37	499	27.9	160	9	US-09-761-120-37	Sequence 37, Appl
38	499	27.9	160	10	US-09-335-325-37	Sequence 37, Appl
39	499	27.9	160	15	US-10-131-241-37	Sequence 37, Appl
40	490	27.4	160	9	US-09-761-120-36	Sequence 36, Appl
41	490	27.4	160	10	US-09-335-325-36	Sequence 36, Appl
42	490	27.4	160	15	US-10-131-241-36	Sequence 36, Appl
43	489	27.4	160	9	US-09-761-120-35	Sequence 35, Appl
44	489	27.4	160	10	US-09-335-325-35	Sequence 35, Appl
45	489	27.4	160	15	US-10-131-241-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1

US-09-761-120-33
 ; Sequence 33, Application US/09761120
 ; Patent No. US20020037847A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Polkman, M. Judah
 ; APPLICANT: O'Reilly, Michael
 ; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmin
 ; FILE REFERENCE: 05940-0151 (43171-252068)
 ; CURRENT APPLICATION NUMBER: US/09761.120
 ; CURRENT FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: 09/309, 821
 ; PRIOR FILING DATE: 1999-03-11
 ; PRIOR APPLICATION NUMBER: 08/866, 735
 ; PRIOR FILING DATE: 1997-05-30
 ; NUMBER OF SEQ ID NOS: 47
 ; SOFTWARE: patentin version 3.0
 ; SEQ ID NO 33
 ; LENGTH: 250
 ; TYPE: PRT
 ; ORGANISM: Bos sp.
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Kringle 1-3
 US-09-761-120-33

Query Match 40.1%; Score 716; DB 9; Length 250;

Best Local Similarity 42.1%; Pred. No. 7.7e-55;

Matches 126; Conservative 43; Mismatches 80; Indels 50; Gaps 7;

QY 8 CYHGGRSYRGISSSTVTGRTCSWSSMIPHWRQRTPENYPNAGLTENYCRNPDSSQK-P 66

Db 1 CKTGQTVRGVTAETKSGVTCQKWSATSPHVPKFSPEKFLAGLEENYCRNPDNDNGP 60

QY 67 WCITTPCVRWYCNLTQCSSETSGVLETFVTVPVPSMEAHSEAAPTQTPVRCYCHGN 126

Db 61 WCYTPDKRYDCIPECED-----KMHCS 87


```

1 FILING DATE: 17-Jun-1999
2 CLASSIFICATION: <Unknown>
3 PRIOR APPLICATION DATA:
4 APPLICATION NUMBER: US/08/612,788
5 FILING DATE: <Unknown>
6 ATTORNEY/AGENT INFORMATION:
7 NAME: Warren, William L.
8 REGISTRATION NUMBER: 36,714
9 REFERENCE/DOCKET NUMBER: 05213-0126
10 TELECOMMUNICATION INFORMATION:
11 TELEPHONE: 404-818-3700
12 TELEFAX: 404-818-3799
13 INFORMATION FOR SEQ ID NO: 32:
14 SEQUENCE CHARACTERISTICS:
15 LENGTH: 250 amino acids
16 TYPE: amino acid
17 STRANDEDNESS: <Unknown>
18 TOPOLOGY: linear
19 MOLECULE TYPE: protein
20 HYPOTHETICAL: NO
21 ANTI-SENSE: NO
22 FRAGMENT TYPE: N-terminal
23 ORIGINAL SOURCE:
24 ORGANISM: Porcine
25 IMMEDIATE SOURCE:
26 CLONE: KI-3
27 SEQUENCE DESCRIPTION: SEQ ID NO: 32:
28 US-09-335-325-32
29
30 Query Match 40.0%; Score 715; DB 10; Length 250;
31 Best Local Similarity 42.5%; Pred. No. 9.4e-55;
32 Matches 127; Conservative 41; Mismatches 81; Indels 50; Gaps 7;
33
34 QY 8 CYHGDRSGYRGSITVTGRTCSWSSMPHWHQRTPEYFNAGLTENYCRNPDSC-QQP 56
35 Db 1 CKTGNKNGYRGTSTKSGVLCQKWSVSSHPKPKYSPKFKPLAGLEENYCRNPDNDEKGP 60
36
37 QY 67 WCYTTPCVRWYCNLTQCSSTESGVLETPVVPVPSMEAHSEAPTEQTPVYRQCYHGN 126
38 Db 61 WCYTTPCVRWYCNLTQCSSTESGVLETPVVPVPSMEAHSEAPTEQTPVYRQCYHGN 126
39
40 QY 127 GQYRGFTFTVTGRTCSWSSMPHWHQRTPEYFNAGLTENYCRNPDADTGPWCFTTD 186
41 Db 88 GEHYEGKISKTSGIEGCSWGSQPHAGLYLPKFPKNNKNNYCRNPDGEPRPWCFTTD 147
42
43 QY 187 PSIRWEYCNLTRCSTEGTVVAPTVIQVPSLGFPEQDQMGNGKGYRGKXATTVTGTP 246
44 Db 148 PKNWEFCIDPRCT-----TTP-----PTSPTQ--CLKGRGNYRGTSVTSAGHT 193
45
46 QY 247 CQEWAAQEPHRSTFTPTGNKWAGLEKNYCRNPDGDIINGPCYTMNPKFLDYCDIPLC 305
47 Db 194 CQRWSAQSPHKHNR-TPENFCKNLEENYCRNPDGE-TAPWCYTTSSEVRWDYCKIPSC 250
48
49 RESULT 6
50 US-10-131-241-32
51 Sequence 32, Application US/10131241
52 Publication No. US20030012792A1
53 GENERAL INFORMATION:
54 APPLICANT: Holaday, John W.
55 TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell
56 TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
57 FILE REFERENCE: 05213-0344 43170-271565
58 CURRENT APPLICATION NUMBER: US/10/131,241
59 CURRENT FILING DATE: 2002-07-22
60 PRIOR APPLICATION NUMBER: US 09/413,049
61 PRIOR FILING DATE: 1999-10-06
62 PRIOR APPLICATION NUMBER: US 09/316,802
63 PRIOR FILING DATE: 1999-05-21
64 PRIOR APPLICATION NUMBER: US 60/086,586
65 PRIOR FILING DATE: 1998-05-22
66 NUMBER OF SEQ ID NOS: 65

```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Porcine
US-10-131-241-32

Query Match          40.0%; Score 715; DB 15; Length 250;
Best Local Similarity 42.3%; Pred. No. 9.4e-54;
Matches 127; Conservative 41; Mismatches 81; Indels 50; Gaps 7;

QY      8 CYHGDGRSYRGISSTVTGTCOSWSSMTPHWHORTPENYPNAGLTENYCRNPDGSG-KQP 66
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1 CKTGNGKNYRGTSKTKSGVICOQWSVSPHYPKYSPEKPLAGLENYCRNPDNDKGP 60
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      67 WCYTTDPCVRWEYCNLTQCSETESGVLETPVVPVPSMEAHSEAAPEQTPTVVRQCYHN 126
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      61 WCYTTDPTREDFDYCDICEED-----ECMHCS 87

QY      127 GOSYRGFTSTVTGTCOSWSSMTPHWHORTPENYPNDGLTMNYCRNPDADTGPWCFTTD 186
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      88 GEHYEGKISKTMGIECQSGWSQSPHAGYLPSEKPNKLNKMYCRNPDGEPWPWCFTTD 147
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      187 PSIRWEYCNLTFRCSDEGTGTVVAPPTVIQVPSLGPPSDQDMFGNGKGYRKKATVTGP 246
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      148 PNKRWEFCIDPRCT-----TPSGPTYQ--CLKGRGENYRGTVSVTASGHT 193
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      247 COENAAQPHRHSTFIPGNTKWAGLEKNYCRNPDGDPGWPCYTMNPKLFYCDIPIC 305
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      194 CQWNSAQSPHKHNR-TPENFPCKNLNENYCRNPDGE-TAPWCYTTDSEVRWDYCKIPSC 250
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
US-10-131-241-61
; Sequence 61, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Fortier, John W.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271585
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-61

Query Match          39.8%; Score 711; DB 15; Length 260;
Best Local Similarity 42.3%; Pred. No. 2.2e-54;
Matches 129; Conservative 40; Mismatches 86; Indels 50; Gaps 7;

QY      5 VDCYHGDGRSYRGISSTVTGTCOSWSSMTPHWHORTPENYPNAGLTENYCRNPDGSGK 64
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      3 LSECKTGNGKNYRGTMSTKNGITCQKWSSTSPHPSFSPATHPSGLENYCRNPDNDP 62
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      65 Q-PWCYTTDPCVRWEYCNLTQCSETESGVLETPVVPVPSMEAHSEAAPEQTPTVVRQCY 123
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      63 QGFWCYTTDPEKRYDYCDICEE-----ECM 89

QY      124 HNGOSYRGFTSTVTGTCOSWSSMTPHWHORTPENYPNDGLTMNYCRNPDADTGPWCFT 183
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      90 HCSGENYDGKISKTMGIECQWDSQSPHAGYIPSPKPNKLNKMYCRNPDRELRPWCF 149
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
QY      184 TTDPSIRWEYCNLTFRCSDEGTGTVVAPPTVIQVPSLGPPSDQDMFGNGKGYRKKATVT 243
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      150 TTDPNKRWELCDIPRCT-----TPSGPTYQ--CLKGTGENYRGVAVTVS 195
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      244 GTPCOENAAQPHRHSTFIPGNTKWAGLEKNYCRNPDGDPGWPCYTMNPKLFYCDIP 303
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      196 GHTCQWNSAQSPHPSF-TPENFPCKNLNENYCRNPDGK-RAPWCYTTNSQVRWEYCKIP 253
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      304 LCASS 308
      1 : : :
Db      254 SCDSS 258
      1 : : :

RESULT 8
US-09-761-120-31
; Sequence 31, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmin
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 31
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Macaca sp.
; NAME/KEY: misc_feature
; OTHER INFORMATION: Kringle 1-3
US-09-761-120-31

Query Match          39.4%; Score 704; DB 9; Length 250;
Best Local Similarity 42.5%; Pred. No. 8.6e-54;
Matches 127; Conservative 40; Mismatches 82; Indels 50; Gaps 7;

QY      8 CYHGDGRSYRGISSTVTGTCOSWSSMTPHWHORTPENYPNAGLTENYCRNPDGSG-KQP 66
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1 CKTGNGKNYRGTMSTKNGITCQKWSSTSPHPSFSPATHPSGLENYCRNPDNDGQGP 60
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      67 WCYTTDPCVRWEYCNLTQCSETESGVLETPVVPVPSMEAHSEAAPEQTPTVVRQCYHN 126
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      61 WCYTTDPEERFDYCDICEED-----ECMHCS 87

QY      127 GOSYRGFTSTVTGTCOSWSSMTPHWHORTPENYPNDGLTMNYCRNPDADTGPWCFTTD 186
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      88 GENYDGKISKTMGIECQWDSQSPHAGYIPSPKPNKLNKMYCRNPDGEPWPWCFTTD 147
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      187 PSIRWEYCNLTFRCSDEGTGTVVAPPTVIQVPSLGPPSDQDMFGNGKGYRKKATVTGTP 246
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      148 PNKRWELCDIPRCT-----TPSGPTYQ--CLKGTGENYRGVAVTVSGHT 193
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      247 COENAAQPHRHSTFIPGNTKWAGLEKNYCRNPDGDPGWPCYTMNPKLFYCDIPIC 305
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      194 CHWSAQTPHPSHNR-TPENFPCKNLNENYCRNPDGE-KAPWCYTTNSQVRWEYCKIPSC 250
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-09-335-325-31
; Sequence 31, Application US/09335325
; Patent No. US2002016471A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; O'Reilly, Micheal
; Cao, Yihai
```

```

; Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Rhesus monkey
; IMMEDIATE SOURCE:
; CLONE: KI-3
; SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-335-325-31

```

```

Query Match      39.4%; Score 704; DB 10; Length 250;
Best Local Similarity 42.5%; Pred. No. 8.6e-54;
Matches 127; Conservative 40; Mismatches 82; Indels 50; Gaps 7;

QY      8 CYHGDGRSYRGISSTVTGRTCSQSSMIPHWHQRTPEPNAGLTENYCRNPDG-KQOP 66
Db      1 CKTGKNTYRGIMTSKTRGITCQKWSSTSPHRTFSPATHPSGEGLEENYCRNPDNDGQGP 60

QY      67 WCYTDPGVWEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAAPTQTPVVRQCYHGN 126
Db      61 WCYTDPGEERFYCDIPECD-----PSSGPTQY--CLAGTGENYRGDVAVTVSGHT 193

QY      127 GQSYRGFTSTVTGRTCSQSSMIPHWHQRTPEPNAGLTENYCRNPDADTGWCFCTTD 186
Db      88 GENDYDKISKTMSGLECCQANDSQSPHAGHYIPSKFFPNKLNKKNYCRNPDGPRPCWCTTD 147

QY      187 PSIRWEYCNLTRCSDEGTGVVAPPTVIQVPSLGPPEQDCMGKGYRGKATTVTGP 246
Db      148 PNKRWELCDIPRCT-----TPP-----PSSGPTQY--CLAGTGENYRGDVAVTVSGHT 193

QY      247 COEWAQAEPRHSTPIPTGNTKAGLENYCRNPDGDIINGPWCYTMNPKLFYDCDIPLC 305
Db      194 CHGWSAQTPHTNR-TPENFCKNLDENYCRNPDGE-KAPWCYITNSQVRWEYCKIPSC 250

RESULT 10
US-10-131-241-31

```

```

; Sequence 31, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 31
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Rhesus monkey
; US-10-131-241-31

Query Match      39.4%; Score 704; DB 15; Length 250;
Best Local Similarity 42.5%; Pred. No. 8.6e-54;
Matches 127; Conservative 40; Mismatches 82; Indels 50; Gaps 7;

QY      8 CYHGDGRSYRGISSTVTGRTCSQSSMIPHWHQRTPEPNAGLTENYCRNPDG-KQOP 66
Db      1 CKTGKNTYRGIMTSKTRGITCQKWSSTSPHRTFSPATHPSGEGLEENYCRNPDNDGQGP 60

QY      67 WCYTDPGVWEYCNLTQCSSTESGVLETPVVPVPSMEAHSEAAPTQTPVVRQCYHGN 126
Db      61 WCYTDPGEERFYCDIPECD-----PSSGPTQY--CLAGTGENYRGDVAVTVSGHT 193

QY      127 GQSYRGFTSTVTGRTCSQSSMIPHWHQRTPEPNAGLTENYCRNPDADTGWCFCTTD 186
Db      88 GENDYDKISKTMSGLECCQANDSQSPHAGHYIPSKFFPNKLNKKNYCRNPDGPRPCWCTTD 147

QY      187 PSIRWEYCNLTRCSDEGTGVVAPPTVIQVPSLGPPEQDCMGKGYRGKATTVTGP 246
Db      148 PNKRWELCDIPRCT-----TPP-----PSSGPTQY--CLAGTGENYRGDVAVTVSGHT 193

QY      247 COEWAQAEPRHSTPIPTGNTKAGLENYCRNPDGDIINGPWCYTMNPKLFYDCDIPLC 305
Db      194 CHGWSAQTPHTNR-TPENFCKNLDENYCRNPDGE-KAPWCYITNSQVRWEYCKIPSC 250

RESULT 11
US-09-761-120-30
; Sequence 30, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Polkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmin
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 30
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Kringle 1-3

```


Db 88 GENTDGRKISTKSLGLEQAWDSQSPHAGYIPKFPNKLKKNYCRNPDRPLPWCFTTD 147
QY 187 PSIRWEYCNLTRCSDEGTGVAPPVTVQVPSLGPSPDQCMFGNGKYGKATTVGTP 246
Db 148 PNKEWELCDIPRCT-----TPP-----PSSGPTYQ--CLAGTGENYRGNAVTVSGHT 193
QY 247 COEWAAOEPHRHSTTFIPGINKWAGLEKNYCRNPDPDINGPWCYTMNPKRLFDYCDIPLC 305
Db 194 CQWSEQTTPHNR--TPENFPCKNLNENYCRNPDPGK--RAPWCHTINSQVRWEYCKIPSC 250

RESULT 14
US-09-761-120-29
; Sequence 29, Application US/097611120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasminogen
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Kringle 1-3
US-09-761-120-29

Query Match 39.2%; Score 700; DB 9; Length 250;
Best Local Similarity 41.1%; Pred. No. 1.9e-53;
Matches 123; Conservative 37; Mismatches 89; Indels 50; Gaps 6;

QY 8 CYHGDGRSYRGISSTVTGRTCSQSSMTPHWHQRTPTENYPNAGLTENYCRNPDSGKO-P 66
Db 1 CKTGIGNGYRGTSRTKSGVACQKGGATFPHPVNPSPSTHPNEGLEENYCRNPDEQGP 60
QY 67 WCYTTPDCVRWEYCNLTQCSSETESGVLETPVVPVPSMEAHSEAAPEQTVPVRCYHGN 126
Db 61 WCYTTPDKRYDYNICEPEE-----PSPTYQCLKGRGENYRGVTSVTSKGT 87
QY 127 GOSYRGFTSTVTGRTCSQSSMTPHWHQRTPTENYPNAGLTENYCRNPDPADTGPWCFTTD 186
Db 88 GEKYEGKISKTMSGLDQAWDSQSPHAGYIPAKFPKSNLKNYCHNPDPGEPWPWCFTTD 147
QY 187 PSIRWEYCNLTRCSDEGTGVAPPVTVQVPSLGPSPDQCMFGNGKYGKATTVGTP 246
Db 148 PTKRWEYCDIPRC-----ATPPPP-----PSPTYQCLKGRGENYRGVTSVTSKGT 193
QY 247 COEWAAOEPHRHSTTFIPGINKWAGLEKNYCRNPDPDINGPWCYTMNPKRLFDYCDIPLC 305
Db 194 CQWSEQTTPHNR--TPENFPCKNLNENYCRNPDPGE--TAPWCYTTPDSQLRWEYCEIPSC 250

RESULT 15
US-09-335-325-29
; Sequence 29, Application US/09335325
; Patent No. US20020164717A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiotensin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K1-3
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-335-325-29

Query Match 39.2%; Score 700; DB 10; Length 250;
Best Local Similarity 41.1%; Pred. No. 1.9e-53;
Matches 123; Conservative 37; Mismatches 89; Indels 50; Gaps 6;

QY 8 CYHGDGRSYRGISSTVTGRTCSQSSMTPHWHQRTPTENYPNAGLTENYCRNPDSGKO-P 66
Db 1 CKTGIGNGYRGTSRTKSGVACQKGGATFPHPVNPSPSTHPNEGLEENYCRNPDEQGP 60
QY 67 WCYTTPDCVRWEYCNLTQCSSETESGVLETPVVPVPSMEAHSEAAPEQTVPVRCYHGN 126
Db 61 WCYTTPDKRYDYNICEPEE-----PSPTYQCLKGRGENYRGVTSVTSKGT 87
QY 127 GOSYRGFTSTVTGRTCSQSSMTPHWHQRTPTENYPNAGLTENYCRNPDPADTGPWCFTTD 186
Db 88 GEKYEGKISKTMSGLDQAWDSQSPHAGYIPAKFPKSNLKNYCHNPDPGEPWPWCFTTD 147
QY 187 PSIRWEYCNLTRCSDEGTGVAPPVTVQVPSLGPSPDQCMFGNGKYGKATTVGTP 246
Db 148 PTKRWEYCDIPRC-----ATPPPP-----PSPTYQCLKGRGENYRGVTSVTSKGT 193
QY 247 COEWAAOEPHRHSTTFIPGINKWAGLEKNYCRNPDPDINGPWCYTMNPKRLFDYCDIPLC 305
Db 194 CQWSEQTTPHNR--TPENFPCKNLNENYCRNPDPGE--TAPWCYTTPDSQLRWEYCEIPSC 250

Search completed: September 29, 2003, 08:18:17
Job time : 349 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:01:08 ; Search time 19 Seconds
(without alignments)
1558.945 Million cell updates/sec

Title: US-10-088-548-2
Perfect score: 1786
Sequence: 1 KSPVQDCVHGGRSYRGIS.....YTMNPKLFEDYCDIPICASS 308

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 160415

Minimum DB seq length: 0
Maximum DB seq length: 308

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.76.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	620	34.7	169	2 A40522	plasmin (EC 3.4.21
2	413	23.1	120	2 E61545	plasmin (EC 3.4.21
3	408	22.8	123	2 C61545	plasmin (EC 3.4.21
4	354	19.8	89	2 A60140	plasmin (EC 3.4.21
5	246.5	13.8	291	2 I38098	t-plasminogen acti
6	88.5	5.0	277	2 I37552	Ox40 homolog - hum
7	88.5	5.0	289	2 AB1075	right origin-bindl
8	87.5	4.9	221	2 A47367	24K androgen-depen
9	86.5	4.8	297	2 JQ1209	attachment protein
10	85.5	4.8	272	2 AB3430	transposase BME114
11	84.5	4.7	241	2 A85231	hypothetical prote
12	84.5	4.7	304	2 T24703	hypothetical prote
13	83.5	4.7	297	2 JQ1208	attachment protein
14	82	4.6	240	2 T33698	hypothetical prote
15	81	4.5	248	2 T19913	hypothetical prote
16	80	4.5	295	2 JC5559	lectin-B - Virgini
17	79.5	4.5	194	2 JC4589	immunoreactive pro
18	79.5	4.5	301	2 D44355	CD44 glycoprotein
19	79	4.4	221	1 S24328	glutathione peroxi
20	78.5	4.4	197	2 B49247	merozoite surface
21	78.5	4.4	287	2 S65765	chitinase (EC 3.2.
22	78	4.4	270	2 AI3110	conserved hypothet
23	78	4.4	270	2 C98176	iob protein [impo
24	77.5	4.3	104	2 S26475	Ig kappa chain v r
25	77.5	4.3	221	1 S24327	glutathione-peroxi
26	77	4.3	284	2 H87524	hypothetical prote
27	76	4.3	166	2 H89044	protein B0238.12 i
28	75.5	4.2	292	1 MGNZ60	major surface glyc
29	75.5	4.2	306	2 T52305	En/Spm-like transp

ALIGNMENTS

RESULT 1

A40522

plasmin (EC 3.4.21.7) precursor - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 28-Feb-1992 #sequence_revision 17-Apr-1993 #text_change 16-Jul-1999

C:Accession: A40522

R:Kanadas, J.J.; Makker, S.P.

J. Biol. Chem. 266, 10825-10829, 1991

A:Title: Identification of the rat Heymann nephritis autoantigen (GP330) as a recept

A:Reference number: A40522; MUID:91250378; PMID:1645711

A:Accession: A40522

A>Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-169 <KAN>

A:Cross-references: GB:M62832; NID:g206215; PIDN:AAA1884.1; PID:g554488

A>Note: the authors translated the codon TCT for residue 76 as Ala

C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homo

C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase

F:34-112/Domain: Kringle homology <KRG>

F:34-112,55-95,83-107/Disulfide Bonds: #status predicted

Query Match	34.7%	Score 620;	DB 2;	Length 169;
Best Local Similarity	61.0%;	Pred. No. 5,1e-39;		
Matches 111;	Conservative 14;	Mismatches 43;	Indels 14;	Gaps 2;
Qy	76	RWEYCNLTQCSSETESGVLETPVWPVPSMEAHSEAAPEOTPVVRCYHGNGCSTRTFS	135	
Db	1	RWEYCEIPSGSGSSVSPDQSDSVLP-----EQTPVQECYOGNGKSYRGTS	47	
Qy	136	TTVTGRTCSQSSMTPEHRTPEYVNDGLTMVCRNPDA--TGPWCFTTDPSTIRNEYC	194	
Db	48	TTNTGKQCSQSSMTPEHRTPEYVNDGLTMVCRNPDA--TGPWCFTTDPSTIRNEYC	107	
Qy	195	NLTRCSDEGVVAPPTVTVQVPSLGPSPQDCMFGNGKGYRGKATVTGTGTCQEWAAE	254	
Db	108	NLKRCSDEGVVAPPTVTVQVPSLGPSPQDCMFGNGKGYRGKATVTGTGTCQEWAAE	167	
Qy	255	PH 256		
Db	168	PH 169		

RESULT 2

E61545

plasmin (EC 3.4.21.7) precursor - dog (fragments)

N:Alternate names: plasminogen

C:Species: Canis lupus familiaris (dog)

C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-May-1995

C:Accession: E61545

R:Schaller, J.; Rickli, E.E.

Enzyme 40, 63-69, 1988

A:Title: Structural aspects of the plasminogen of various species.

whey acidic protei
 glutathione peroxi
 leucocyte common a
 hypothetical prote
 secretory protein
 myo-inositol catab
 Ig kappa chain v r
 Ig kappa chain v r
 Ig kappa chain v r
 hypothetical prote
 iolB protein [impo
 major surface glyc
 coocidiosis-relate
 hypothetical prote
 Ig kappa chain v r
 Ig kappa chain v r

A:Reference number: A61545; MUID:89005015; PMID:3168975
A:Accession: E61545
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-120 <SCH>
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: hydrolase; serine proteinase
F:37-114/Domain: kringle homology <KR4>

Query Match 23.1%; Score 413; DB 2; Length 120;
Best Local Similarity 74.7%; Pred. No. 7.1e-24;
Matches 65; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

QY 119 VRCYHGQSGYRGFTTGTGTCQSWSMTPHRRHQTPENYNDGLTMYCRNPDAQT 178
Db 34 VQECYHGQSGYRGFTTGTGTCQSWSMTPHRRHQTPENYNDGLTMYCRNPDAQT 93

QY 179 GPCWFTDPSIRWEYCNLTRCSDTEGT 205
Db 94 SPMCYTDPFSRWEYCNLTRCSDTEGT 120

RESULT 3
C61545
plasmin (EC 3.4.21.7) precursor - goat (fragments)
N:Alternate names: plasminogen
C:Species: Capra aegagrus hircus (domestic goat)
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-May-1995
C:Accession: C61545
R:Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A:Title: Structural aspects of the plasminogen of various species.
A:Reference number: A61545; MUID:89005015; PMID:3168975
A:Accession: C61545
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-123 <SCH>
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: hydrolase; serine proteinase
F:41-118/Domain: kringle homology <KR4>

Query Match 22.8%; Score 408; DB 2; Length 123;
Best Local Similarity 73.9%; Pred. No. 1.7e-23;
Matches 68; Conservative 9; Mismatches 13; Indels 2; Gaps 1;

QY 110 AAPTEQTPVRCYHGQSGYRGFTTGTGTCQSWSMTPHRRHQTPENYNDGLTMY 169
Db 31 AAKCEE--AQDCYHGQSGYRGFTTGTGTCQSWSMTPHRRHQTPENYNDGLTMY 88

QY 170 YCRNPDAQTGPWCFTDPSIRWEYCNLTRCSD 201
Db 89 YCRNPDAKSPWCYTTDPFRWERCNLKGCSE 120

RESULT 4
A60140
plasmin (EC 3.4.21.7) precursor - chicken (fragment)
N:Alternate names: plasminogen
C:Species: Gallus gallus (chicken)
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 16-Jul-1999
C:Accession: A60140
R:Gyenes, M.; Pathy, L.
Biochim. Biophys. Acta 832, 326-330, 1985
A:Title: The kringle 4 domain of chicken plasminogen.
A:Reference number: A60140; MUID:86077796; PMID:4074753
A:Accession: A60140
A:Molecule type: protein
A:Residues: 1-89 <GY>
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F:6-83/Domain: kringle homology <KR>
F:6-83,27-66,55-78/Disulfide bonds: #status predicted
F:39/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 19.8%; Score 354; DB 2; Length 89;
Best Local Similarity 67.9%; Pred. No. 1.2e-19;
Matches 55; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

QY 121 QCYHGQSGYRGFTTGTGTCQSWSMTPHRRHQTPENYNDGLTMYCRNPDAQTGP 180
Db 5 ECYQNGVSYRGFTTGTGTCQSWSMTPHRRHQTPENYNDGLTMYCRNPDAQTGP 64

QY 181 WCFTTDPFSIRWEYCNLTRCSD 201
Db 65 WCFTTDPFSIRWEYCNLTRCSD 85

RESULT 5
I38098
t-plasminogen activator precursor, inactive endothelial splice form - human
N:Alternate names: tissue plasminogen activator
C:Species: Homo sapiens (man)
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
C:Accession: I38098; S01678
R:Siebert, P.D.; Fong, K.
Nucleic Acids Res. 18, 1086, 1990
A:Title: Variant tissue type plasminogen activator (PLAT) cDNA obtained from human en
A:Reference number: I38098; MUID:90191218; PMID:1969145
A:Accession: I38098
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-291 <SIE>
A:Cross-references: EMBL:X13097; NID:g35282; PIDN:CAA31489.1; PID:g35283
C:Comment: For the main splice form, see PIR:UKHUT. This form probably does not have
C:Genetics:
A:Gene: GDB:PIAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat
C:Keywords: alternative splicing; fibrinolysis; glycoprotein; kringle
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-32/Domain: propeptide #status predicted <PRO>
F:33-291/Product: t-plasminogen activator, inactive endothelial splice form #status p
F:41-78/Domain: fibronectin type I repeat homology <IFA>
F:86-119/Domain: EGF homology <EGF>
F:127-208/Domain: kringle homology #status atypical <KR1>
F:215-291/Domain: kringle homology #status atypical <KR2>
F:41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203/Disulfide bonds: #status p

Query Match 13.8%; Score 246.5; DB 2; Length 291;
Best Local Similarity 26.2%; Pred. No. 3.7e-11;
Matches 80; Conservative 34; Mismatches 110; Indels 81; Gaps 14;

QY 14 RSYRGI---SSTVTGTCQSWSMTPHRRHQTPENYNDGLTMYCRNPDSGKQPCYT 70
Db 35 RSYVICRDEKTMQYQHQSWLRPVL--SNRVEYCWGSGRAQCHSVYKSCSEPCFN 93

QY 71 TDCVRWEXCN--LQCCSETESGVLETPVVPVPSMEAHSEAAPEQTPVVRQCYHGNGQ 128
Db 94 GGTCQALYFSDVQCQPEGFAG-----KCCEIDTRAT-----CYEDQGI 133

QY 129 SYRGTSTVTGTCQSWSMTPHRRHQ---RTPENYNDGLTMY-----NYCRNPDAQTGP 180
Db 134 SYRGTSTVTAESGAECTNWNSSALACNAYSGRRP-----DAIRLGLGNHYCRNPDRDQSP 188

QY 181 WCFTTDP--SIRWEYCNLTRCSDTEGTVPVAPPYVPSLPSPSEDCQMGNGKGYRGKA 239
Db 189 WCIVFRAGKYSSEFCSTPACSE-----GNSDCYFGNGSAIRGTHS 228

QY 240 TTVTGTGTCQSW-----AAQEPHRRHSTFIPGTNKGWGLEK--NYCRNPDSGNGPWC 288
Db 229 LTESGASCLPWNSSMILGKVTQAQNSAQA-----LGLGKHYCRT--GRSVSSPAT 278
QY 289 YTMNP 293
:|

Db 279 ASMRP 283

RESULT 6
OX40 homolog - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 11-Jan-2000
C:Accession: I37552
R:Latza, U.; Durkop, H.; Schnittger, S.; Ringeling, J.; Eitelbach, F.; Hummel, M.; Fonat
Eur. J. Immunol. 24, 677-683, 1994
A:Title: The human OX40 homolog: cDNA structure, expression and chromosomal assignment
A:Reference number: I37552; MUID:94170844; PMID:7510240
A:Accession: I37552
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-277 <RES>
A:Cross-references: EMBL:X75962; NID:g472957; PIDN:CAA53576.1; PID:g472958
C:Superfamily: CD27 antigen; NGF receptor repeat homology

Query Match 5.0%; Score 88.5; DB 2; Length 277;
Best Local Similarity 19.0%; Pred. No. 17;
Matches 51; Conservative 31; Mismatches 89; Indels 97; Gaps 12;

QY 18 GISSTVTGRTG-C-QSWSSMIPHHQRTPENYPNAGLTENYCRNPDGSKQPCW----- 68
DB 20 GLGSTVTGLHCVGDTFSDRCCHCECPGN----GWSRCRSRONTVCRP-CGPGFYND 74

QY 69 -YTTDCVRWEXCNLTQCSSETSGVLETPTW-----PVPSMEAHSEAAPTQTPVV 119
DB 75 VVSSKCKPCTWCNLRSSERKQLCTATQDTCRCRAGTQPLDSYKPGVDCAPCPP---- 130

QY 120 RQCVHNGQSYRGTSTVTGRTQCSWSMT-PHRHQRTPENYPNDGLTMNYCRNPDAAT 178
DB 131 -----GHFSPG-DNQACKPWTNCTLAGKHTLQFASNSSDAI----- 165

QY 179 GPWCFTTDPDSIRWEYCNLTGCSDEGTGTVVAPPTVQVPSLGGPSQDCMFGNGKGYRGKK 238
DB 166 -----CHDRD-----PPATQQTQGGPP----- 183

QY 239 ATTVTGPPCOEW--AAQEPHRSHTFIPG 264
DB 184 ARPITVQTEAWPTSQGSTPRVEVPG 211

RESULT 7
AB1075
right origin-binding protein [imported] - Salmonella enterica subsp. enterica serovar Ty
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AB1075
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Farry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AB1075
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD03417.1; PID:ig16505685; GSPDB:GN00176
C:Genetics:
A:Gene: STY4933

Query Match 5.0%; Score 88.5; DB 2; Length 289;
Best Local Similarity 22.7%; Pred. No. 18;
Matches 48; Conservative 18; Mismatches 60; Indels 85; Gaps 10;

QY 20 SSTVTGTCSSSM-----IPHHQRTPENYPNAGLTENYCRNPDGSKQPV 67

Db 97 SQTALYRRSPWSAFGRPRLRLGEFTVPEHOFVLTEDPLLGVTQSY----- 146

QY 68 CYTTDCVRWEXCNLTQCSSETSGVL-----EPTTVPEV-----PSMEAHSE- 109
DB 147 -----CSLEQISDFRHEMRVQFWHDFLGHSPSTIPPEVLYGLNETRPSMEKDDQ 194

QY 110 -----AAPEQT-----PVVRQ-----CYHNGOSYRGTFSTTVTGRTQCSWS 148
DB 195 EVFTTALFOEAGDGVSAPVLLQGEVMTYEGLTGVQ-DFILTVYG-TCMPMLN 252

QY 149 MTPHRHQRTPENYPNDGLTMNYCRNPDAATG 179
DB 253 LTRKGDIERYPSE-----DKTGTG 273

RESULT 8
A47367
24K androgen-dependent glutathione peroxidase-like protein arMEP24 - mouse
N:Alternate names: androgen-regulated secretory protein arMEP24
C:Species: Mus musculus (house mouse)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 11-Jun-1999
C:Accession: A47367; S13725
R:Ghyssels, N.B.; Dufau, I.; Lareyre, J.J.; Rigaudiere, N.; Mattei, M.G.; Dufau,
Mol. Endocrinol. 7, 258-272, 1993
A:Title: Structural organization and regulation of the gene for the androgen-dependent
A:Reference number: A47367; MUID:93225968; PMID:8469239
A:Accession: A47367
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-221 <GHY>
A:Experimental source: BALB/c
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBI:129287, NCBI:129289)
R:Ghyssels, N.B.; Dufau, J.P.
Nucleic Acids Res. 18, 7144, 1990
A:Title: A mouse cDNA sequence for epididymal androgen-regulated proteins related to
A:Reference number: S13725; MUID:91088305; PMID:2263479
A:Accession: S13725
A:Molecule type: mRNA
A:Residues: 47-52, 'H', 54-102, 'Q', 104-169, 'S', 171-221 <GH2>
A:Cross-references: EMBL:X53780; NID:g50023; PIDN:CAA37796.1; PID:g50024
C:Superfamily: glutathione peroxidase

Query Match 4.9%; Score 87.5; DB 2; Length 221;
Best Local Similarity 26.4%; Pred. No. 16;
Matches 29; Conservative 9; Mismatches 39; Indels 33; Gaps 4;

QY 193 YCNLTGCSDEGTGTVVAPPTVQVPSLGGPSQDCMFGNGKGYRGKATTVTGTPCOEWAA 252
DB 72 YGLT-----IQYELNALQEDLAPFG-----LVILGFPNCFGK 106

QY 253 QPHERHSTFTPG---TNKWLGLKYNCRNPDGDLNGPWCTMNPRLFDY 299
DB 107 QEPGDNLEILLGLKYVRPGKGLPNQLFAKGDVNGE-----NEQKIFTF 151

RESULT 9
JQ1209
attachment protein - human respiratory syncytial virus (strain RSB6614)
N:Alternate names: G protein
C:Species: human respiratory syncytial virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Feb-1997
C:Accession: JQ1209
R:Cane, P.A.; Matthews, D.A.; Pringle, C.R.
J. Gen. Virol. 72, 2091-2096, 1991
A:Title: Identification of variable domains of the attachment (G) protein of subgroup
A:Reference number: JQ1204; MUID:91374005; PMID:1895054
A:Accession: JQ1209
A:Molecule type: mRNA
A:Residues: 1-297 <CAN>
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract
er children and adults.
C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C;Keywords: glycoprotein; transmembrane protein
F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 4.8%; Score 86.5; DB 2; Length 297;
Best Local Similarity 23.0%; Pred. No. 26;
Matches 57; Conservative 26; Mismatches 88; Indels 77; Gaps 13;
QY 66 PWCYTDPCVRWEYCNLTQCSETSGVLETFETVPVPSNEAHSEAAPEQTTPVRCYHG 125
Db 88 PTLQNPQLGHSFNL---SETTSQATTPALT-TPSAES-----TPGSTTVK----- 132
QY 126 NGQSYRGTSITVTGTCQSSMTPHRHQRTPENYPNDGL---TMNYCRNPADATGPMC 182
Db 133 -----TKNTTIT---QLQP-SKPTTKQHQKPPKPNHHEFEVFNFCSCSNNPTC 182
QY 183 F-----TTDPSIRWEYCNLTQCSDEGTGVVAPTVTVQVSLGPPSF 223
Db 183 WAJCKRIPNKKGKTKTKTKPKTKIK-----TTKKDLKPQTKPKKEVLTT-----KPT 232
QY 224 QDCMFGNGKGYRGKKAAT-VTGTPCQEWAAQEPHRSFTFPGTKWAGLEKNYCRNPDDG 282
Db 233 KPINTTKNIRITLLTNTTNGNP--EYTSQKETHST-----SPGN 273
QY 283 INGPWCYT 290
Db 274 PPSQVYT 281

RESULT 10
AB3430
transposase BMEI1424 [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: AB3430
R;DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Majer, C.; Ios, T.; Ivanova,
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A;Reference number: AD3252; PMID:11756688
A;Accession: AB3430
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-272 <KUR>
A;Cross-References: GB:AE008917; PIDN:AA152605.1; PID:g17983424; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI1424
A;Map position: 1

Query Match 4.8%; Score 85.5; DB 2; Length 272;
Best Local Similarity 21.8%; Pred. No. 28;
Matches 62; Conservative 29; Mismatches 77; Indels 117; Gaps 18;
QY 30 QSSSMTPHWHQRTPENYNAGLTENY-----CRNPSGKQPCWYTTDCV- 75
Db 10 QQAVIAP-----LLPNQGAHRITDDRVISGIIHLRSGCRWQDC---PACIGPPTVY 62
QY 76 ----RWE-----YCNLTQCSDETSQVLETFETVPVPSNEAHSEAA----- 111
Db 63 NRPFWAKGIWRLEALVQPTDRIHMDSTTA-----KAHRAAGGKGDAEAGR 116
QY 112 -----PTEQPPVVRQCYHGQSYRGVGT-----FSTTV-----TGRT 142
Db 117 SRGGRSTKIAHVDSC--GRPVALRITPGQGDAPITVILLDVSQEVVHSERIEETGVT 174
QY 143 QSSSMTPHWHQRTPENYPNDGLT-----MNYCRNPADTG-----PW-----CFTDPS 198
Db 175 TPQSFRIAPNEH---PQECPTDAIASRNRNGRSGFRQADGAGACTPLMGLAENCFLSDRT 231
QY 189 IR--WEYCNLTQCSDEGTGVVAPTVTVQVSLGP-----PSEQ 224
Db 232 VODIWP-----GGDDGLSVAPH--OPPSNGFGHCRITDTPSQ 267

RESULT 11
AB5231
hypothetical protein AT4g20350 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C;Accession: AB5231
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: AB5001; MUID:20083488; PMID:10617198
A;Accession: AB5231
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-241 <STO>
A;Cross-References: GB:NC_001268; NID:g7268830; PIDN:CAB79035.1; GSPDB:GN00140
C;Genetics:
A;Gene: AT4g20350
A;Map position: 4

Query Match 4.7%; Score 84.5; DB 2; Length 241;
Best Local Similarity 21.8%; Pred. No. 29;
Matches 42; Conservative 16; Mismatches 70; Indels 65; Gaps 9;
QY 95 TPTVVPVPSNEAHSEAAPEQTTPVRCYHGNGQSTRGTFTTGTTCQSSMTPHRH 154
Db 11 TPTVVPVPGF-----ITDEQTQLLNHIYGASGSKW-----TLNRRLLQNWGMV-HEK 59
QY 155 QRTPENYP-----NDGLTMNYCRNPADATGPGWCTTDPDSIRWEY 193
Db 60 GLVPQELFPWLTKTAEIHESGLFPSSAINHVLNEY--HPDQGIMP----- 104
QY 194 CNLTRCSDTEGTVPVAPPVTVQVPSLGPPSEQD---CMFNGKGYRGKKAATVTGTCOE 249
Db 105 -----HODGPAIYFP--VVALISLGSPVYMDPTPHLRSGDGYISKDQ-----SPCAE 150
QY 250 WAAQEPHRSFTFI 262
Db 151 SCAPERDSFVLL 163

RESULT 12
T24703
hypothetical protein T08G3.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T24703
R;Lloyd, C.
submitted to the EMBL Data Library, December 1996
A;Reference number: Z19926
A;Accession: T24703
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-304 <WIL>
A;Cross-References: EMBL:283238; PIDN:CAB05796.1; GSPDB:GN00023; CESP:T08G3.6
A;Experimental source: clone T08G3
C;Genetics:
A;Gene: CESP:T08G3.6
A;Map position: 5
A;Introns: 86/2

Query Match 4.7%; Score 84.5; DB 2; Length 304;
Best Local Similarity 19.1%; Pred. No. 37;
Matches 53; Conservative 33; Mismatches 108; Indels 83; Gaps 14;
QY 87 ETESGVLETFVVPV-----PMEAHSEAAPEQ--TP-----VTRQC 122
Db 26 EINFSEIFKTRIVPLMSVQMACGCVSGAPRRNALQNAQKICTPGSACRQISLIYRVC 85
QY 123 YHGNGQSYRGVTFSTTVT-----GRTC-----QSWSSMTPHRHQRTPEN--YPND 164
Db 86 ---XGDTYRILCNKTCVPAQRTCCIPVYPMINGSQCGCFPPDFPDPSSESFCPCPKG 142

QY 165 GLTMN---YCRNPADTGPW-----CFTDPSIRWEYCNLTFRCSDTGTVVAP 209
 Db 143 GFWSWWSAYFR--DGEKKAWSTRRLCTEEAGCCTDPTGTTIETSTACFCRKLVDVSSLV 200
 QY 210 PTVIQVPSIGPP-SEQDCMGNGKGYRGKATVTGTPCOEWAAQPHRHSTFI-----PG 264
 Db 201 KTNLGTYLTPVYNDEDCASQVM---ESQKSAVGECNEYSWKYYIWNNAVRYKPG 256
 QY 265 TNNKAGLEKNYCRNPDDINGPWCYTMNPKLFYCD 301
 Db 257 RSDYVGIRISNCE-----AAHEKKIFLYCD 281

RESULT 13
 JQ1208
 Attachment protein - human respiratory syncytial virus (strain RS6256)
 N:Alternate names: G protein
 C:Species: human respiratory syncytial virus
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Feb-1997
 C:Accession: JQ1208
 R:Canine, P.A.; Matthews, D.A.; Pringle, C.R.
 J. Gen. Virol. 72, 2091-2096, 1991
 A:Title: Identification of variable domains of the attachment (G) protein of subgroup A
 A:Reference number: JQ1204; MUID:91374005; PMID:1895054
 A:Accession: JQ1208
 A:Molecule type: mRNA
 A:Residues: 1-297 <CAN>
 C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract ill-
 children and adults.
 C:Superfamily: Respiratory syncytial virus major surface glycoprotein G
 C:Keywords: glycoprotein; transmembrane protein
 F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 4.7%; Score 83.5; DB 2; Length 297;
 Best Local Similarity 22.6%; Pred. No. 43;
 Matches 56; Conservative 26; Mismatches 89; Indels 77; Gaps 13;

QY 66 PWCYTIDPCVWEYCNLTQSETESGVLETPVVPVPSMEAHSEAAFTQTPTVPRQCYHG 125
 Db 88 PYYLQNQLGISFNL---SETTS-----QPTTPAPTPTPS---ABSTPQSTTVK----- 132

QY 126 NGQSVRGTFSTVTRGTCQSWSSMTPHRHQRTPENYPNDGL---TMNYCRNPADTGPWC 182
 Db 133 -----TKNITIT--QIQP-SKPTTKQKQKPPKPNNDPHEVFNVPSCISNNPTC 182

QY 183 F-----TTDPSIRWEYCNLTFRCSDTGTVVAPTVIQVPSIGPSE 223
 Db 183 WAICKRIPNKKRGKKTTKTKPTIK-----TTKDLKPQTKPKREVLT-----KPT 232

QY 224 QDCMFGNGKGYRGKAT--VTGTPCOEWAAQPHRHSTFIPGTNKNAGLEKNYCRNPDD 282
 Db 233 KPTINTRTNRTILLTNTGNP--EYTSQKEILHST-----SPEGN 273

QY 283 INGFWCYT 290
 Db 274 PPSQVYT 281

RESULT 14
 T33698
 hypothetical protein F49f1.7 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 08-Sep-2000
 C:Accession: T33698
 R:Miller, N.; Wamsley, P.
 submitted to the EMBL Data Library, October 1998
 A:Description: The sequence of C. elegans cosmid F49f1.
 A:Reference number: Z21389
 A:Accession: T33698
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-240 <MIL>
 A:Cross-references: EMBL:AF100656; PIDN:AA668946.1; GSPDB:GN00022; CESP:F49f1.7

A:Experimental source: strain Bristol N2; clone F49f1
 C:Genetics:
 A:Gene: CESP:F49f1.7
 A:Map position: 4
 A:Introns: 42/2; 79/2; 116/2; 209/2
 C:Superfamily: Caenorhabditis elegans hypothetical protein F49f1.6

Query Match 4.6%; Score 82; DB 2; Length 240;
 Best Local Similarity 21.1%; Pred. No. 44;
 Matches 44; Conservative 17; Mismatches 74; Indels 74; Gaps 9;

QY 28 TCQSWSSMIPHWHORT-----PENYPNAGLTENYCRNPDSGKQPCWCTTDCVWRWEYCNL 82
 Db 57 TCSIDNSDCEKFEHMCVMPYRQCPKS-----CGICDSPATWCM-----HWPECPK 105

QY 83 TC-CSETE-----SGVLETPVVPVPSMEAHSEAAFTQT 116
 Db 106 YQVPCSETERIKCPRSCDAGNPSTKEASVTITNGHLQIGTTRRLQATQITKSTETTETT 165

QY 117 PVVRQCYHNGQSVRGTFSTVTRGTCQSWSSMTPHRHQRTPENYPNDGLTMNYCRNDA 176
 Db 166 ETK-----AIKTKTKTETTKATTPSTTVTKTKP-----KPECT 204

QY 177 DTGPPW-----FTTD---PSIRWEYC 194
 Db 205 DSSPNCITWAKNGFCTNTFYPPKREYC 233

RESULT 15
 T19913
 hypothetical protein C43f9.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T19913
 R:Mortimore, B.
 submitted to the EMBL Data Library, November 1996
 A:Reference number: Z19195
 A:Accession: T19913
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-248 <MIL>
 A:Cross-references: EMBL:Z82262; PIDN:CAB05152.1; GSPDB:GN00022; CESP:C43f9.5
 A:Experimental source: clone C43f9
 C:Genetics:
 A:Gene: CESP:C43f9.5
 A:Map position: 4
 A:Introns: 47/1; 78/3; 118/1; 199/3

Query Match 4.5%; Score 81; DB 2; Length 248;
 Best Local Similarity 23.8%; Pred. No. 55;
 Matches 59; Conservative 19; Mismatches 114; Indels 56; Gaps 15;

QY 13 GRSYRGISS---TVVGTQCSWSSIP---HWHTPTPENYPNAGLTENYCRNPDSGKQFW 67
 Db 20 GLTELGIKSKRQTHICGTYPNQFYSPCEYWTSTQVFN-----PYTCANGGRKIGVG 72

QY 68 CYTTDPCVW---EYCNLTQSETESGVLETPVVPVPSMEAHSEAAFTQTPTVVRQCYH 124
 Db 73 CYTYNQCTPYAANSVC-LNNCCCTNPVVTTRAPITPTT-----TFYSALAYCYN 122

QY 125 GNQSVRGTFST-TVTGTCQSWSSMTPHRHQRTPENYPND--GL-TMNYCRNPADTGP 180
 Db 123 GQRTQVRCVTSVDCAAAGQTCMNGICCT-----TTGNEYTSCGGLPAISAC-----GTGQ 172

QY 181 WC---FTTDPSTWEYCNLTFRCSDTGTVVAPTVIQVPSIGPSEQDCMGNGKGYRGK 238
 Db 173 TCGSEFCTSS---NYCCEQYGTAG-----LCSNGCPTGYSONSNTNGY--CC 217

QY 239 ATTVGTGP 246
 Db 218 ATCAGRP 225

Wed Oct 1 15:58:32 2003

us-10-088-548-2.sep29.rpr

Page 6

Search completed: September 29, 2003, 08:04:08
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:01:08 ; Search time 11 Seconds
(without alignments)
1316.748 Million cell updates/sec

Title: US-10-088-548-2
Perfect score: 1786
Sequence: 1 KSPVVQCYHGDGRSTRGIS.....YTMNPKLFYCDIPCLASS 308

Scoring table: BLASUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 65706

Minimum DB seq length: 0

Maximum DB seq length: 308

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	620	34.7	169	1 PLMN_RAT	Q01177 rattus norv
2	88.5	5.0	277	1 TNRA_HUMAN	P43489 homo sapien
3	86.5	4.8	297	1 VGLG_HRSV7	P27026 human respi
4	84.5	4.7	221	1 GSHE_MOUSE	P21765 mus musculu
5	83.5	4.7	297	1 VGLG_HRSV6	P27025 human respi
6	79	4.4	221	1 GSHE_RAT	P30710 rattus norv
7	77.5	4.3	221	1 VGLG_HRSV8	P28714 macaca fasc
8	75.5	4.2	292	1 VGLG_HRSV8	P23041 human respi
9	75	4.2	127	1 WAP_RABIT	P09412 oryctolagus
10	75	4.2	221	1 GSHE_RAT	O64625 rattus norv
11	74.5	4.2	219	1 GSHE_PIG	O18994 sus scrofa
12	74	4.1	224	1 XP4_XENLA	O00223 xenopus lae
13	74	4.1	271	1 IOLB_BACSU	P42413 bacillus su
14	73.5	4.1	292	1 VGLG_HRSV1	P20896 human respi
15	73	4.1	265	1 KNHL_CANGA	O74884 candida gla
16	73	4.1	300	1 TR6B_HUMAN	O95407 homo sapien
17	72.5	4.1	217	1 SGS3_DROSI	P13729 drosophila
18	72.5	4.1	221	1 GSHE_HUMAN	O75175 homo sapien
19	72.5	4.1	269	1 WBP1_HUMAN	O96427 homo sapien
20	72	4.0	263	1 SGS3_DROYA	P13728 drosophila
21	72	4.0	272	1 TNRA_MOUSE	P47741 mus musculu
22	72	4.0	297	1 VGLG_HRSV4	P27023 human respi
23	71.5	4.0	162	1 HXA9_CAVPO	P51783 cavia porce
24	71.5	4.0	221	1 GSHE_CANFA	O46607 canis fami
25	71	4.0	226	1 GSHP_HUMAN	P22352 homo sapien
26	71	4.0	227	1 CTG5_HUMAN	O9br10 homo sapien
27	71	4.0	308	1 HMB1_STRPU	P13545 strongyloce
28	70.5	3.9	169	1 GPO_SYNY3	P74250 synecocyst
29	70.5	3.9	261	1 Y612_SYNY3	P72581 synecocyst
30	70	3.9	241	1 WAP8_HUMAN	O81uao homo sapien
31	69.5	3.9	283	1 HXA9_FUGRU	O42506 fuqu rubrip
32	69.5	3.9	307	1 HMB1_TRIGR	P09080 tritonastes
33	69	3.9	199	1 EOST_ACTEQ	P81439 actinia equ

RESULT 1	PLMN_RAT	STANDARD;	PRT;	169 AA.
34	69	3.9	237	1 IPDE_DICDI
35	69	3.9	272	1 ERG_MOUSE
36	68.5	3.8	278	1 VGLL_HCMV2
37	68	3.8	275	1 UL11_HCMV2
38	67.5	3.8	226	1 GSHP_RAT
39	67.5	3.8	278	1 VGLL_HCMV1
40	67.5	3.8	304	1 YQOB_CAREL
41	67	3.8	154	1 GP41_BPSP1
42	67	3.8	238	1 ET3_HUMAN
43	66.5	3.7	258	1 PRIA_LENED
44	66.5	3.7	278	1 VGLL_HCMV4
45	66.5	3.7	278	1 VGLL_HCMV5

ALIGNMENTS

AC Q01177; DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasminogen (EC 3.4.21.7) (Fragment).
GN PLG
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Kanalas J.J., Makker S.P.;
RA MEDLINE=91250378; PubMed=1645711;
RT receptor site for plasminogen."
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
J. Biol. Chem. 266:10825-10829(1991).
RL -|- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
LAMININ AND VON WILLEBRAND FACTOR.
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
higher selectivity than trypsin. Converts fibrin into soluble
products.
CC -|- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC -|- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -|- SIMILARITY: Contains 5 kringle domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@sib-sib.ch).

DR EMBL: M62832; AAA41884.1; -
DR PIR: A40522; A40522.
DR HSP: P00747; IPMK.
DR MEROPS: S01.233; -
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003966; Prothrombin.
DR InterPro: IPR001254; Ser protease_Try.
DR Pfam: PF00051; kringle; 2.

P22549 dictyosteli
P81270 mus musculu
Q68668 human cytom
P16721 human cytom
P23764 rattus norv
Q68667 human cytom
Q09300 caenorhabdi
Q48397 bacterioph
P14138 homo sapien
Q01200 leishmania
Q68670 human cytom
Q68671 human cytom

```
DR PRINTS; PRO0018; KRINGLE.
DR PRINTS; PRO1505; PROTHROMBIN.
DR PRODOM; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE.1; 1.
DR PROSITE; PS00070; KRINGLE.2; 2.
DR PROSITE; PS02040; TRYPSIN_DOM; PARTIAL.
DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW Hydrolyase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Repeat.
FT NON_TER 1
FT DOMAIN 1 10 KRINGLE 3 (BY SIMILARITY).
FT DOMAIN 34 112 KRINGLE 4 (BY SIMILARITY).
FT DOMAIN 139 >169 KRINGLE 5 (BY SIMILARITY).
FT DISULFID 34 112 BY SIMILARITY.
FT DISULFID 55 95 BY SIMILARITY.
FT DISULFID 83 107 BY SIMILARITY.
FT NON_TER 169 169
SQ SEQUENCE 169 AA; 18401 MW; 77A54214C49D010C CRC64;

Query Match 34.7%; Score 620; DB 1; Length 169;
Best Local Similarity 61.0%; Pred. No. 1.9e-42;
Matches 111; Conservative 14; Mismatches 43; Indels 14; Gaps 2;

QY 76 RWECNLTQCSSETSGVLETPVVPSPMEAHSEAPRTQTTPVRCQYHGNGSQSYRGTF 135
D 1 RWECETPSGSGVSPQSDSVLP-----SQTPVQECYQNGKSYRGTS 47
D 136 TTVGTRCQSNMTPHRTPTNPNDGLTMYCRNPDAD-TGPGWFTTDPDSIRWEYC 194
D 48 TITGKCKQSWMTPHSEKSKPANFPDPSGLENNYCRNPDNDQPGWFTTDPDSIRWEYC 107
QY 195 NLTCSRSTEGTVAPTVTVQVPSGLPPSEDCMGKGYRGKATTVTGPCQEWAAQE 254
D 108 NLKCRSETGGVAESAIQVPSAPGTSYEDCMFGNGKEYRGKTAVTAAGPCQEWAAQE 167
QY 255 PH 256
D 168 PH 169

RESULT 2
TNR4_HUMAN STANDARD; PRT; 277 AA.
AC P43489; Q13663;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 4 precursor (OX40L
DE glycoprotein 1 receptor) (TAX-transcriptionally activated
DE glycoprotein 1 receptor) (CD134 antigen).
GN TNFRSF4 OR TXGP1L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94170844; PubMed=7510240;
RA Latza U., Duerkop H., Schmittger S., Ringeling J., Eitelbach F.,
RA Hummel M., Fonatsch C., Stein H.;
RT "The human OX40 homolog: cDNA structure, expression and chromosomal
RT assignment of the ACT35 antigen.";
RL Eur. J. Immunol. 24:677-683(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95219871; PubMed=7704935;
RA Baum P.R., Gayle R.B. III, Ramsdell F., Srinivasan S., Sorensen R.A.,
RA Watson M.L., Seldin M.F., Clifford K.N., Grabstein K., Alderson M.R.;
RT "Identification of OX40 ligand and preliminary characterization of
RT its activities on OX40 receptor.";
RL Circ. Shock 44:30-34(1994).
```

[illegible]

```

DR EMBL; M68896; AAA37729.1; -
DR EMBL; X53780; CAA37796.1; -
DR PIR; A47367; A47367.
DR HSP; P00435; IGP1.
DR MGD; MGI:104886; Gpx5.
DR InterPro; IPR000889; Glut_peroxidase.
DR Pfam; PF00255; GSHPx; 1.
DR PRINTS; PR01011; GLUTPEROXDASE.
DR PROSITE; PS00460; GLUTATHIONE_PEROXID_1; 1.
DR PROSITE; PS00763; GLUTATHIONE_PEROXID_2; 1.
DR OXIDOREDUCTASE; Peroxidase; Signal.
KW SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 221 EPIDIDYMAL SECRETORY GLUTATHIONE
FT ACT_SITE 73 73 BY SIMILARITY.
FT CONFLICT 53 53 D -> H (IN REF. 2).
SQ SEQUENCE 221 AA; DESF8BD6CD22D6F9 CRC64;

Query Match 4.7%; Score 84.5; DB 1; Length 221;
Best Local Similarity 25.5%; Pred No. 5.6;
Matches 28; Conservative 10; Mismatches 39; Indels 33; Gaps 4;

QY 193 YCNLRCSDEGTGVAPPVTVVQVPSLGPSPQDCMFGNGKGYRGKATVGTGTCQSWAA 252
DQ 72 YCGLT-----IQVELNALQEDLKPFQ-----IWLGFPCNQFGK 106
QY 253 QEPHRSHTFIP---TNKWLGLEKNYCRNPDGDLNGFWCYTMNPKLFDY 299
DQ 107 QEPGNLEILPLGYRPGKGLFNLQFLXAGDVNGE-----NEQKIFTF 151

RESULT 5
VGLG_HRSV6 STANDARD; PRT; 297 AA.
AC P27025;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (strain rsb256).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11256;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
DR PIR; J01208; J01208.
DR InterPro; IPR000925; Glycoprot_G.
DR Pfam; PF00802; Glycoprotein_G; 1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66 POTENTIAL.
FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 297 AA; 32708 MW; 6781756C38B64A80 CRC64;

Query Match 4.7%; Score 83.5; DB 1; Length 297;

```

```

Best Local Similarity 22.6%; Pred. No. 9.3;
Matches 56; Conservative 26; Mismatches 89; Indels 77; Gaps 13;

QY 66 PWCTTDPQVRWEYCNLTQCSETSGVLETPVVPVPSMEAHSEAPTEQFPVVRQCYHG 125
DQ 88 PLYLTQNPQLGISFNSNL---SETTS---QPTTPAPTTPS---AESTPQSTTVK----- 132
QY 126 NGQSYRGFTSTTVTGRTQSWSSMTPHRHQRTPENYPNDGL---TMNYCRNPADATGPMC 182
DQ 133 -----TKNTTTT---QIOP-SKPTTKQRQKNPKPNKPNDFEFVNFVPCSGICNNPTC 182
QY 183 F-----TTDPSIRWEYCNLTQCSDEGTGVAPPVTVVQVPSLGPSPSE 223
DQ 183 WAICKRIPNKKPGKKTITTKPKPTIK-----TTKKDLKPQTKPKVLT-----KPT 232
QY 224 QDCMFGNGKGYRGKATV-VGTGTCQSWAAQEPHRSHTFIPGTNKGWGLEKNYCRNPDGD 282
DQ 233 KPTITRTNIRTLTTLTNTTGNP---EYTSQKELHST-----SPEGN 273
QY 283 INGPWCYT 290
DQ 274 PSPQVTV 281

RESULT 6
GSHE_RAT STANDARD; PRT; 221 AA.
AC P30710;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Epididymal secretory glutathione peroxidase precursor (EC 1.1.1.1.9)
DE (Epididymis-specific glutathione peroxidase-like protein) (EGLP).
GN GPX5
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Wistar; TISSUE=Epididymis;
RX MEDLINE=922559957; PubMed=1386734;
RA Perry A.C.F., Jones R., Niang L.S.P., Jackson R.M., Hall L.;
RT "Genetic evidence for an androgen-regulated epididymal secretory
RT glutathione peroxidase whose transcript does not contain a
RT selenocysteine codon.";
RL Biochem J. 285:863-870(1992).
CC -1- FUNCTION: Protects cells and enzymes from oxidative damage, by
CC catalyzing the reduction of hydrogen peroxide, lipid peroxides and
CC organic hydroperoxide, by glutathione. May constitute a
CC glutathione peroxidase-like protective system against peroxide
CC damage in sperm membrane lipids.
CC -1- CATALYTIC ACTIVITY: 2 glutathione + H(2)O(2) = oxidized
CC glutathione + 2 H(2)O.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: EPIDIDYMIS.
CC -1- SIMILARITY: Belongs to the glutathione peroxidase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; X62404; CAA44274.1; -.
DR PIR; S24528; S24528.
DR HSP; P00435; IGP1.
DR InterPro; IPR000889; Glut_peroxidase.
DR Pfam; PF00255; GSHPx; 1.
DR PRINTS; PR01011; GLUTPROXDASE.
DR PROSITE; PS00460; GLUTATHIONE_PEROXID_1; 1.

```


DR PROSITE; PS00763; GLUTATHIONE_PEROXID_2; 1.
KW Oxidoreductase; Peroxidase; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 221 EPIDIDYMAL SECRETORY GLUTATHIONE
FT ACT_SITE 73 73 BY SIMILARITY.
FT PEROXIDASE.
SQ SEQUENCE 221 AA; 25384 MW; 0D9D3FAC9F12D16 CRC64;
Query Match 4.4%; Score 79; DB 1; Length 221;
Best Local Similarity 25.0%; Pred. No. 15;
Matches 24; Conservative 9; Mismatches 35; Indels 28; Gaps 3;
QY 193 YCNLTRCSSTEGTVVAPTVIQVPSLGPPSEDCMFGNGKGYRGKATVGTGTCQEWAA 252
DB 72 YCGLT-----IQYELNALQDLKQFG-----LVILGPPCNGFGK 106
QY 253 QEPHRSHTFIPG---TNKWAGLEKYNCRNPDGIDNG 285
DB 107 QEPGDNYEILPGLKYVRGKGLFNFLFAKGDVNG 142
RESULT 7
GSHE.MACFA
ID GSHE.MACFA STANDARD; PRT; 221 AA.
AC P28714;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Epididymal secretory glutathione peroxidase precursor (PC 1.11.1.9)
DE (Epididymis-specific glutathione peroxidase-like protein) (EGLP).
GN GPX5.
OS Macaca fascicularis (Crab eating macaque) (Cynomolpus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Epididymis;
RX MEDLINE=92359957; PubMed=1386734;
RA Perry A.C.F., Jones R., Niang L.S.P., Jackson R.M., Hall L.;
RT "Genetic evidence for an androgen-regulated epididymal secretory
glutathione peroxidase whose transcript does not contain a
selenocysteine codon."
RL Biochem. J. 285:863-870(1992).
CC -!- FUNCTION: Protects cells and enzymes from oxidative damage, by
catalyzing the reduction of hydrogen peroxide, lipid peroxides and
organic hydroperoxide, by glutathione. May constitute a
glutathione peroxidase-like protective system against peroxide
damage in sperm membrane lipids.
CC -!- CATALYTIC ACTIVITY: 2 glutathione + H(2)O(2) = oxidized
glutathione + 2 H(2)O.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: EPIDIDYMIS.
CC -!- SIMILARITY: Belongs to the glutathione peroxidase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

CC EMBL; X62403; CAA44273.1; -.
DR PIR; S24327; S24327.
DR HSP; P00435; IGF1.
DR InterPro; IPR000889; Glut_peroxidase.
DR Pfam; PF00255; GSHPx; 1.
DR PRINTS; PR01011; GLUTPROXDASE.
DR PROSITE; PS00460; GLUTATHIONE_PEROXID_1; 1.
DR PROSITE; PS00763; GLUTATHIONE_PEROXID_2; 1.
KW Oxidoreductase; Peroxidase; Signal.

FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 221 EPIDIDYMAL SECRETORY GLUTATHIONE
FT ACT_SITE 73 73 BY SIMILARITY.
FT PEROXIDASE.
SQ SEQUENCE 221 AA; 25212 MW; 32AE912ED6C73D93 CRC64;
Query Match 4.3%; Score 77.5; DB 1; Length 221;
Best Local Similarity 23.6%; Pred. No. 20;
Matches 26; Conservative 11; Mismatches 40; Indels 33; Gaps 4;
QY 193 YCNLTRCSSTEGTVVAPTVIQVPSLGPPSEDCMFGNGKGYRGKATVGTGTCQEWAA 252
DB 72 YCGLT-----AQYELNALQDLKQFG-----LVILGPPCNGFGK 106
QY 253 QEPHRSHTFIPG---TNKWAGLEKYNCRNPDGIDNGPWCYTMNPKLFY 299
DB 107 QEPGDNKEILGLKYVRGPGGFVNFQLFKGDVNGE-----KEQKVFSE 151
RESULT 8
VGLG_HRSV8
ID VGLG_HRSV8 STANDARD; PRT; 292 AA.
AC P23041;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (subgroup B / strain G/60).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11258;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=90357765; PubMed=1697126;
RA Sullender W.M., Anderson K., Wertz G.W.;
RT "The respiratory syncytial virus subgroup B attachment glycoprotein:
RT analysis of sequence, expression from a recombinant vector, and
RT evaluation as an immunogen against homologous and heterologous
RT subgroup virus challenge."
RL Virology 178:195-203(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374595; PubMed=1895391;
RA Sullender W.M., Wufson M.M., Anderson L.J., Wertz G.W.;
RT "Genetic diversity of the attachment protein of subgroup B
RT respiratory syncytial viruses."
RL J. Virol. 65:5425-5434(1991).
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
HEMAGGLUTININATING ACTIVITIES.
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

CC EMBL; M56633; AAA47413.1; -.
DR EMBL; M73545; AAA47408.1; -.
DR PIR; A37077; MGN260.
DR InterPro; IPR000925; Glycoprot_G.
DR Pfam; PF00802; Glycoprotein_G; 1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66 POTENTIAL.
FT DOMAIN 67 292 EXTRACELLULAR (POTENTIAL).

```

PTR; S01286; S01286.
HSP; 046635; ICJH.
InterPro: IPR002221; WAP.
Pfam: PF00095; wap; 2.
ProDom: PD001224; WAP_C; 1.
SMART: SM00217; WAP; 2.
PROSITE; PS00317; 4_DISULFIDE_CORE; 2.
Milk; Whey; Protease inhibitor; Repeat; Phosphorylation; Signal.
FT SIGNAL 1 19
CHAIN 20 127
DOMAIN 28 68
DOMAIN 74 121
MOD_RES 237 237
SEQUENCE 127 AA; 13526 MW; 59DAD5F8FD2E2087 CRC64;

Query Match 4.2%; Score 75; DB 1; Length 127;
Best Local Similarity 26.2%; Pred.No.17;
Matches 17; Conservative 9; Mismatches 31; Indels 8; Gaps 2;

QY 57 CRNPDGPGWCYITDPCVRWEYCNLTQCSFETPTVTVVPTSMSEAHSEAAFTPTQT 116
DB 32 CPEPSSSEETICLSDNCLGSTVC----CPSAGSGSCTPIIVPTPK----AGRCPWQA 83

QY 117 PVVRQ 121
DB 84 PMLSQ 88

RESULT 10
ID GSHY_RAT STANDARD; PRT; 221 AA.
AC Q64625;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Glutathione peroxidase RX2D1 precursor (EC 1.11.1.9) (Odorant-
DE metabolizing protein RX2D1).
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RX [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Olfactory epithelium;
RX MEDLINE=92031476; PubMed=1931961;
RT Dear T.N., Campbell K., Rabbitts T.H.;
RA "Molecular cloning of putative odorant-binding and odorant-
RT metabolizing proteins."
RL Biochemistry 30:10376-10382(1991).
CC -!- CATALYTIC ACTIVITY: 2 glutathione + H(2)O(2) = oxidized
CC glutathione + 2 H(2)O.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE BOWMAN'S GLANDS.
CC -!- SIMILARITY: Belongs to the glutathione peroxidase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M76733; AAA42094.1; -.
CC PIR; B40464; B40464.
CC HSP; P00435; IGPI.
CC InterPro: IPR000889; Glut_peroxidase.
CC Pfam; PF00255; GSHPx; 1.
CC PRINTS; PF01011; GLUTPROXDASE.
CC PROSITE; PS00460; GLUTATHIONE_PEROXID_1; 1.
CC PROSITE; PS00763; GLUTATHIONE_PEROXID_2; 1.
CC oxidoreductase; Peroxidase; Signal.
CC FT SIGNAL 1 21
CC POTENTIAL.

```

FT	CHAIN	22	221	GLUTATHIONE PEROXIDASE RY2D1.
FT	ACT_SITE	73	73	BY SIMILARITY.
SQ	SEQUENCE	221 AA;	24961 MW; 92749EAF6A3EF48C CRC64;	
 Query Match 4.2%; Score 75; DB 1; Length 221; Best local similarity 25.0%; Pred. NO. 31; Matches 23; Conservative 13; Mismatches 32; Indels 24; Gaps 3;				
QY	227 MFNGSGYGRKKAT-----	TVTGTPCOEWAAQEPHHRHTFPG---TNK 267		
	: : : : :	: : : : : : : : : : :		
Db	65 LFVNASFCGLRATYPRLNTLQELRPFNVSVLGPFPCNFGKQEPGNSEILLGLKYVRP 124			
QY	268 WAGLENYCNPDGDINGPCWYNPRKLFDY 299			
	: : : : :	: : : : : : : : : : :		
Db	125 GGGFVFNFOLFKEGVNDG-----NQKVFSF 151			
 RESULT 11 GSHE_PIG STANDARD; PRG: 219 AA. ID GSHE_PIG STANDARD; PRG: 219 AA. AC OI8994; DT 15-JUL-1999 (Rel. 38, Created) DT 15-JUL-1999 (Rel. 38, Last sequence update) DT 30-MAY-2000 (Rel. 39, Last annotation update) DE Epidiymal secretory glutathione peroxidase precursor (BC 1.11.1.9) DE (Epidiymis-specific glutathione peroxidase-like protein) (ESLP). GN GPX5. OS Sus scrofa (pig). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OX Mammalia; Theria; Cetartiodactyla; Suina; Suidae; Sus. NCBI_TaxID=9823; [1] RN SEQUENCE FROM N.A., AND SEQUENCE OF 22-46. RC TISSUE-Epididymis; RX MEDLINE=97415456; PubMed=9271255; OKamura N., Iwaki Y., Hiramoto S., Tamba M., Bannai S., Sugita Y., Syntin P., Dacheux F., Dacheux J.L.; RA "Molecular cloning and characterization of the epididymis-specific RT glutathione peroxidase-like protein secreted in the porcine RT epididymal fluid". RL Biochim. Biophys. Acta 1336:99-109(1997). CC -! FUNCTION: MAY CONSTITUTE A GLUTATHIONINE PEROXIDASE-LIKE CC PROTECTIVE SYSTEM AGAINST PEROXIDE DAMAGE IN SPERM MEMBRANE CC LIPIDS. SINCE THE PURIFIED PORCINE ENZYME HAS VERY LITTLE ACTIVITY CC TOWARDS HYDROGEN PEROXIDE OR ORGANIC HYDROPEROXIDES THE PROTECTIVE CC EFFECT IS NOT LIKELY TO BE EXERTED BY ITS ENZYMATIC ACTIVITY. CC INSTEAD, MAY PROTECT SPERM FROM PREMATURE ACROSOME REACTION IN THE CC EPIDIYMS BY BINDING TO LIPID PEROXIDES, WHICH MIGHT OTHERWISE CC INTERACT WITH PHOSPHOLIPASE A2 AND INDUCE THE ACROSOME REACTION. CC -! CATALYTIC ACTIVITY: 2 glutathione + H(2)O(2) = oxidized CC glutathione + 2 H(2)O. CC -! SUBUNIT: Homotetramer. CC -! SUBCELLULAR LOCATION: Secreted. CC -! TISSUE SPECIFICITY: PROXIMAL CAPUT EPIDIYMS. CC -! SIMILARITY: Belongs to the glutathione peroxidase family. This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
DR	EMBL; D37916; BAA22149.1; -			
DR	HSSP; P00435; 1GP1.			
DR	InterPro; IPR000889; Glut_peroxidase.			
DR	Pfam; PF00255; GSHPx; 1.			
DR	PRINTS; PR01011; GLUTPROXDASE.			
DR	PROSITE; PS00460; GLUTATHIONE_PEROXID_1; 1.			
DR	PROSITE; PS00763; GLUTATHIONE_PEROXID_2; 1.			
KW	Oxidoreductase; Peroxidase; Signal.			
FT	SIGNAL	1	21	

```

FT DISULFID 96 113 BY SIMILARITY.
FT DISULFID 125 152 BY SIMILARITY.
FT DISULFID 136 161 BY SIMILARITY.
FT DISULFID 146 163 BY SIMILARITY.
FT DISULFID 175 201 BY SIMILARITY.
FT DISULFID 185 200 BY SIMILARITY.
FT DISULFID 195 212 BY SIMILARITY.
FT CARBOHYD 104 104 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 224 AA; 24815 MW; 6AEO5B3429EFC635 CRC64;

Query Match 4.1%; Score 74; DB 1; Length 224;
Best Local Similarity 22.2%; Pred. No. 38;
Matches 38; Conservative 25; Mismatches 70; Indels 38; Gaps 10;

QY 42 RTPENYFAGLTENYC-----RNPDSGKQPCWYTTDPCVRWEYCNLTQCSFESGV 92
Db 32 KSRDNCPPGIGSPDCVKKGCCFDSDPDPS---IWCYTP-----WKFEF-TICNPAEPKA 82
QY 93 LETFTVVPVPSMAHSEAAPTQT--PVVRCYHGNGQSYRGTFSTVTGRTCSWSMTP 151
Db 83 RVNGCYGITSQCDKKGCCFNITPNNVCYQPIEAVE-----RDC---SAVEP 130
QY 152 HRHQ--TPENYNDGLTMVCRNPADATGFCWFTTDPSSIRWEYCNLTFRCS 200
Db 131 KRVNCGPPGVSPDECINKGCGFNSDVGVPWCF--KPEIKKE---LIQCA 176

RESULT 13
ID IOLB_BACSU STANDARD; PRT; 271 AA.
AC P42413;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE IOLB protein.
GN IOLB OR E83B.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RA Azevedo V., Bessieres P., Boller A., Borchert S.,
RA Borriass R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Brington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Neone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,

```

```

RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
RN [3]
RP SEQUENCE OF 1-71 FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE=9603926; PubMed=7584049;
RA Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;
RT "Cloning and sequencing of a 36-kb region of the Bacillus subtilis
RT genome between the gnt and iol operons."
RL DNA Res. 2:61-69(1995).
CC -!- PATHWAY: Myo-inositol catabolism.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D14399; BAA03291.1; -
CC EMBL; Z99124; CAB16011.1; -
CC EMBL; AB005554; BAA21610.1; -
CC PIR; B69645; B69645.
CC Subtilisin; BGI1118; iolB.
CC Complete proteome.
KW SEQUENCE 271 AA; 30770 MW; 4322AE4C18DF21F3 CRC64;

Query Match 4.1%; Score 74; DB 1; Length 271;
Best Local Similarity 20.3%; Pred. No. 47;
Matches 42; Conservative 19; Mismatches 64; Indels 82; Gaps 10;

QY 101 VPSMEAHSEAAPEQTPTV-----VRQCY-----HGMGQSYR 131
Db 76 IGRSEYFPERKPTDSVIISNDRAFEITAVSDARVALCYSPSEKOLPKLIAEDNGIEHR 135
QY 132 GTSTTIVTGT-----CQSSMTPHRHORTPENYDNG-LTNY 170
Db 136 GQFSNKRRTVNIPLDPSANSLLVVEVYDTSNWSYFPHKHDQ--DNLPEESFLEET 193
QY 171 CRNPADATG---PWCFTTDPSSIRWEYCNLTGSDTEGTWVAPTVIOVPS---LGPPE 223
Db 194 YHLDPCQGFVFORVYTDTSI-----DETWTGNNVIVPAGTHPVGVDP- 240
QY 224 QDCMFGNGKGYRGKATTVTGTQCEW 250
Db 241 -----GYTSYLVNMG-PTRKW 257

RESULT 14
VGLG_HRSV1
ID VGLG_HRSV1 STANDARD; PRT; 292 AA.
AC P20896;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-NOV-1991 (Rel. 20, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (subgroup B / strain 18537).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11251;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=87289657; PubMed=2441388;
RA Johnson P.R., Spriggs M.K., Olmsted R.A., Collins P.L.;

```

RT "The G glycoprotein of human respiratory syncytial viruses of
 RT subgroups A and B: extensive sequence divergence between
 RT antigenically related proteins";
 RL J. Bacteriol. 180:5020-5029(1998).
 CC -!- FUNCTION: INVOLVED IN CELL WALL BETA(1->6) GLUCAN SYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: SECRETED, PROBABLY FOUND AT CELL SURFACE (BY
 CC SIMILARITY).
 CC -!- PTM: O-GLYCOSYLATED (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE KRE9 / KNH1 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M17213; AAA47412.1; -
 DR PIR: B32703; MGNE18.
 DR InterPro: IPR000925; Glycoprot G.
 DR Pfam: PF00802; Glycoprotein_G; 1.
 KW Transmembrane; Glycoprotein.
 FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 38 66 POTENTIAL.
 FT DOMAIN 67 292 EXTRACELLULAR (POTENTIAL).
 FT CARBOHYD 81 81 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 292 AA; 32306 MW; BC8C59F69CA7AFC2 CRC64;
 Query Match 4.1%; Score 73.5; DB 1; Length 292;
 Best Local Similarity 20.6%; Pred. No. 56;
 Matches 49; Conservative 36; Mismatches 98; Indels 53; Gaps 9;
 QY 19 ISSTVTGTCQSSSMIPHHQRTPEPNAGLLENYCRNPDGKOPWCYTDFCVRW 78
 DB 69 VLTITVTVQIKN-----HTKNI-STYLTQVPERVNSKQP---TITSPIITN 114
 QY 79 YCNLTQCSSTES-----GVLETFVVPVPSMEAHSEAPTEQTQPVVRQCYH----- 124
 DB 115 SATISPTKSETHHTTAQTGRITTTQINKPSTRSKNPKKP---KDDYHFEVNF 170
 QY 125 -----GNCQSVRGTFSTVIGR-----TCQSSSMTPHHQRTPEPNAGLLEN 169
 DB 171 VPCISGNNOLCKSICKTIPSNKPKKPTIKPTNKPTTKTKRDPKPAKPKKEIITN 230
 QY 170 YCRNP-----DADTGFWCFTTDFPSIRWEYCNLTGCSDTGTVVAPPTVIQVPSLGPPS 222
 DB 231 PAKKPKTKTERTDSISQSTVLDITIPKY-TIQQCSLHSTSTENTPSSTQIPTASEPS 287
 RESULT 15
 KNH1_CANGA
 ID KNH1_CANGA STANDARD; PRT; 265 AA.
 AC O74684;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Cell wall synthesis protein KNH1 precursor.
 GN KNH1.
 OS Candida glabrata (Yeast) (Torulopsis glabrata).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5478;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 2001;
 RX MEDLINE=98422452; PubMed=9748432;
 RA Nagahashi S., Iussier M., Bussey H.;
 RT "Isolation of Candida glabrata homologs of the Saccharomyces

RT cerevisiae KRE9 and KNH1 genes and their involvement in cell wall
 RT beta-1,6-glucan synthesis";
 RL J. Bacteriol. 180:5020-5029(1998).
 CC -!- FUNCTION: INVOLVED IN CELL WALL BETA(1->6) GLUCAN SYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: SECRETED, PROBABLY FOUND AT CELL SURFACE (BY
 CC SIMILARITY).
 CC -!- PTM: O-GLYCOSYLATED (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE KRE9 / KNH1 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF064252; AAC64009.1; -
 KW Glycoprotein; Cell wall; Signal.
 FT SIGNAL 1 14 POTENTIAL.
 FT CHAIN 15 265 CELL WALL SYNTHESIS PROTEIN KNH1.
 SQ SEQUENCE 265 AA; 29046 MW; 6A13BCD6BC0AE2BC CRC64;
 Query Match 4.1%; Score 73; DB 1; Length 265;
 Best Local Similarity 20.1%; Pred. No. 55;
 Matches 57; Conservative 38; Mismatches 88; Indels 100; Gaps 16;
 QY 52 LLENYCRNPDGKOP--WCYTDFPCVRWEYCNLT--OCSETSGVLETFVVPVPSMEAH 107
 DB 23 LDQSFEDASGGTAKIPTQWLFTPTNPQDFTSLTFLSCG-----PNYKIEAF 70
 QY 108 -----SEAPTEQTQPVVRQCYHGQSGYRGTFSTVTGTCQSSSMTPHHQRTPEPN 161
 DB 71 KVIGKLSDIGTDFEAEVSQSGVANGYIYVOITAA----- 106
 QY 162 PNDGLTMYCRNPDAITGPWCETTDPSIRWEYCNLT--RCSDEGTGVVAPTV---IQVP 216
 DB 107 -TDGTYIY-----SPRKLTCMTGSKLPDLT-LITAPPTPETRVTTG 147
 QY 217 SLGPPSEQ---DCMFG--NGKG---YEGKATITVGTFCQEWAAQ----- 253
 DB 148 DLGATIDSKSFDIPYGEONGKAKFAPMOTQPGKITAT--TWSRRYATSAVSFTSLTA 204
 QY 254 EPHRSTFIPGTNKGWAGLEKNYC-RNPDGDLINGPWCYTMNPK 295
 DB 205 TPVQHTLTFGWSYIISDYNYAPAPPSPDNGM---YDPKK 244
 Search completed: September 29, 2003, 08:02:14
 Job time : 14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:01:08 ; Search time 36 Seconds
(without alignments)
2207.783 Million cell updates/sec

Title: US-10-088-548-2

Perfect score: 1786

Sequence: 1 KSPVVQDCYHGDSRGYS.....YTMNPKLFYCDIFLCASS 308

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 330525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 513101

Minimum DB seq length: 0

Maximum DB seq length: 308

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTRMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	644	36.1	113	4 Q9UIR5	Q9uir5 homo sapien
2	634	35.5	113	4 Q9UIP6	Q9uir6 homo sapien
3	550.5	30.8	215	13 Q42341	Q42341 gallus gall
4	544	30.5	95	4 Q8N696	Q8n696 homo sapien
5	498	27.9	132	4 Q16609	Q16609 homo sapien
6	489	27.4	113	4 Q9UIR7	Q9uir7 homo sapien
7	481.5	27.0	290	4 Q02935	Q02935 homo sapien
8	477.5	26.7	296	4 Q14519	Q14519 homo sapien
9	467.5	26.2	145	6 Q28911	Q28911 macaca fasc
10	466	26.1	285	4 Q8TCE2	Q8tce2 homo sapien
11	445	24.9	105	4 Q9UIR8	Q9uir8 homo sapien
12	261.5	14.6	109	6 Q9N1B8	Q9n1b8 ovis aries
13	260	14.6	209	11 Q8BS17	Q8bs17 mus musculus
14	249.5	14.0	208	4 Q8BYN0	Q8byn0 homo sapien
15	249.5	14.0	210	4 Q13494	Q13494 homo sapien
16	248	13.9	53	4 Q9UMI2	Q9um12 homo sapien

17	243.5	13.6	211	11 Q55027	O55027 mus musculus
18	241	13.5	60	4 Q9UKJ7	Q9ukj7 homo sapien
19	231.5	13.0	90	4 Q8NG20	Q8ng20 homo sapien
20	225	12.6	202	13 Q90675	Q90675 gallus gall
21	210.5	11.8	75	6 Q9EGN9	Q9egn9 bos taurus
22	168.5	9.4	111	6 Q77688	Q77688 oryctolagus
23	165.5	9.3	157	6 Q9TVA8	Q9tva8 bos taurus
24	159	8.9	231	11 Q8C6L2	Q8c6l2 mus musculus
25	152	8.5	154	4 Q96SE8	Q96se8 homo sapien
26	141	7.9	25	4 Q9UD88	Q9ud88 homo sapien
27	117	6.6	263	4 Q00318	Q00318 homo sapien
28	117	6.6	263	4 Q96FE7	Q96fe7 homo sapien
29	117	6.6	263	4 Q8NCJ9	Q8ncj9 homo sapien
30	114	6.4	162	11 Q8C4E2	Q8c4e2 mus musculus
31	114	6.4	286	4 Q13209	Q13209 homo sapien
32	95.5	5.3	214	6 Q9XT70	Q9xt70 oryctolagus
33	91	5.1	261	5 Q9VAX5	Q9vax5 drosophila
34	90.5	5.1	268	10 Q42830	Q42830 phytophthor
35	89.5	5.0	289	16 Q8ZJ07	Q8zj07 salmonella
36	88.5	5.0	289	16 Q8Z0T3	Q8z0t3 salmonella
37	87.5	4.9	223	4 Q96RT0	Q96rt0 homo sapien
38	87	4.9	249	10 Q24102	Q24102 medicago tr
39	86.5	4.8	223	4 Q96BH3	Q96bh3 homo sapien
40	86.5	4.8	295	12 Q86361	Q86361 respiratory
41	85.5	4.8	272	16 Q8YF06	Q8yfu6 brucella me
42	85	4.8	234	11 Q9CU03	Q9cu03 mus musculus
43	84.5	4.7	220	4 Q9H4C8	Q9h4c8 homo sapien
44	84.5	4.7	221	11 Q8CDQ5	Q8cdq5 mus musculus
45	84.5	4.7	241	10 Q9SUP1	Q9sup1 arabidopsis

ALIGNMENTS

RESULT 1
Q9UIR5 PRELIMINARY; PRT; 113 AA.
AC Q9UIR5;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21181705; PubMed=11285247;
RA Ogorelkova M., Kraft H.G., Ehnholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -|- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158663; AAF03680.1; -
DR HSSP; P00747; 1PMK.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12685 MW; F3D65681D9B5253A CRC64;
Query Match 36.1%; Score 644; DB 4; Length 113;
Best Local Similarity 99.1%; Pred. No. 1.2e-49;

Matches 112; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 112 PTEQIPVRCYHGNGQSYRGFTSTVTGRTCSWSMTPHRRHQTPTNPGLTMYTC 171
 DB 1 PTEQIPVRCYHGNGQSYRGFTSTVTGRTCSWSMTPHRRHQTPTNPGLTMYTC 60

QY 172 RNPADGTPWCFTTDPISIRWEYCNLTCSOTECTVAPPTVIVQVSLGPPSEQ 224
 DB 61 RNPADGTPWCFTTDPISIRWEYCNLTCSOTECTVAPPTVIVQVSLGPPSEQ 113

RESULT 2

Q9UR6 PRELIMINARY; PRT; 113 AA.

AC Q9UR6; PRT; 113 AA.

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Apolipoprotein(a) (Fragment).

GN APOA.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=21181705; PubMed=11285247;

RT "Single nucleotide polymorphisms in exons of the apo(a) kringle IV
 types 5 to 10 domain affect Lp(a) plasma concentrations and have
 different patterns in Africans and Caucasians.";

RL Hum. Mol. Genet. 10:815-824(2001).

CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR EMBL; AF158661; AAF03679.1; -

DR EMBL; AF158660; AAF03679.1; JOINED.

DR HSP; P00747; 2PK4.

DR InterPro; IPR000001; Kringle.

DR Pfam; PF00051; kringle; 1.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 1.

DR SMART; SM00130; KR; 1.

DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.

KW Glycoprotein; Kringle; Lipoprotein.

FT NON_TER 1

FT NON_TER 113

SQ SEQUENCE 113 AA; 12697 MW; 51D4461D9C66312E CRC64;

Query Match 35.5%; Score 634; DB 4; Length 113;
 Best Local Similarity 100.0%; Pred. No. 9.2e-49;
 Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSPVQDCYHGNGSYRGISSTVTGRTCSWSMTPHRRHQTPTNPGLTMYTCRNP 60
 DB 4 KSPVQDCYHGNGSYRGISSTVTGRTCSWSMTPHRRHQTPTNPGLTMYTCRNP 63

QY 61 DSGKQPCWYTTDPCVRWEYCNLTQCSSETESGVLETPVVPVPSMEHSEA 110
 DB 64 DSGKQPCWYTTDPCVRWEYCNLTQCSSETESGVLETPVVPVPSMEHSEA 113

RESULT 3

O42341 PRELIMINARY; PRT; 215 AA.

AC O42341;

DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE HGF alpha-chain (Fragment).

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Heart;

RA IsoKawa K., Yahagi N., Honma J., Tanaka M., Murakami K., Yokoyama H.,
 Toda Y.;

RT "The expression of hepatocyte growth factor mRNA is temporally
 relevant to cardiac endothelial-mesenchymal transformation.";

RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 3 KRINGLE DOMAINS.

DR EMBL; D63779; BAA23643.1; -

DR HSP; P00747; IKN.

DR InterPro; IPR000001; Kringle.

DR Pfam; PF00051; kringle; 3.

DR PRINTS; PS00018; KRINGLE.

DR ProDom; PD000395; Kringle; 3.

DR SMART; SM00130; KR; 2.

DR PROSITE; PS00021; KRINGLE_1; 3.

DR PROSITE; PS00070; KRINGLE_2; 3.

KW Glycoprotein; Kringle.

FT NON_TER 1

FT NON_TER 215

SQ SEQUENCE 215 AA; 24949 MW; 55E67AB52DAF316 CRC64;

Query Match 30.8%; Score 550.5; DB 13; Length 215;
 Best Local Similarity 40.9%; Pred. No. 4.9e-41;
 Matches 104; Conservative 24; Mismatches 81; Indels 45; Gaps 6;

QY 44 PENTPAGLTENYCRNPDSGKQPCWYTTDPCVRWEYCNLTQCSSETESGVLETPVVPVPS 103
 DB 2 PERPDGFDNDYCRNPDSGKQPCWYTTDPCVRWEYCNLTQCSSETESGVLETPVVPVPS 50

QY 104 MEAHSEAPTEQTPVYVRCYHGNGQSYRGFTSTVTGRTCSWSMTPHRRHQTPTNP 163
 DB 51 -----TEAVAETTCIOGQEGYRGTVTWTWISGICQQRWDQFPQHNTIPNFKC 101

QY 164 DGLTMYCRNPADGTPWCFTTDPISIRWEYCNLTQCSSETESGVLETPVVPVPS 223
 DB 102 KDLRENYCRNPDSGSPFCFTDPNIRIGYCS-----QIFKCDYSNE 143

QY 224 QDCMFGNGKGYRGKATVTGTTPQEW--AAQEPHRRHTFI--PGTNKAGLEKNTCRNP 279
 DB 144 QDCYRGNGKSYNGNLSTKTFGLTCTWKNIEDLRHRIQIFREPDVSK---LKNYCRNP 200

QY 280 DGDINGPWCYTMP 293
 DB 201 DDDSGPWCYTDDP 214

RESULT 4

Q8N696 PRELIMINARY; PRT; 95 AA.

AC Q8N696;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Antiangiogenic AK38 protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RA Dou D.;

RT "Antiangiogenic domain of apolipoprotein(a).";
 Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR EMBL; AY039748; AAK74187.1; -

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003966; Prothrombin.

DR Pfam; PF00051; kringle; 1.

DR PRINTS; PR00018; KRINGLE.

```

DR PRINTS; PRO1505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR Glycoprotein; Kringle.
SQ SEQUENCE 95 AA; 10553 MW; 31EDC9DB38A372DF CRC64;

Query Match 30.5%; Score 544; DB 4; Length 95;
Best Local Similarity 98.9%; Pred. No. 7.1e-41;
Matches 92; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 215 VPSLGPSPQDCMFGNGKRGKATVTGTCQWAAQEPHRSFIPGTHKWAIGLKN 274
DB 3 VPSLGPSPQDCMFGNGKRGKATVTGTCQWAAQEPHRSFIPGTHKWAIGLKN 62
QY 275 YCNRPDGLNGWCVYTMNPKRLDYCDIPLCA 307
DB 53 YCNRPDGLNGWCVYTMNPKRLDYCDIPLCA 95

RESULT 5
Q16609
ID Q16609 PRELIMINARY; PRT; 132 AA.
AC Q16609;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE (APOARGC).
GN APOARGC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=95268939; PubMed=7749817;
RA Ojorekova M., Kraft H.G., Enholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT types 6 to 10 domain affect lip(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158659; AAF03678.1; -.
DR HSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12815 MW; 4F80ADF8708548CB CRC64;

Query Match 27.4%; Score 489; DB 4; Length 113;
Best Local Similarity 78.0%; Pred. No. 6.5e-36;
Matches 85; Conservative 5; Mismatches 19; Indels 0; Gaps 0;

QY 1 KSPVQDCYHGDGSRYSRISSTVTGTCQSWSSMIPHHORTPENYPNAGLTENCRNP 60
DB 4 QSPVQDCYHGDGSRYSRISSTVTGTCQSWSSMIPHHORTPENYPNAGLTENCRNP 63
QY 61 DSGKQPCWCTTDDPCRWECYCNLTQCSSTESGVLETPVTVVPSMEAHSE 109
DB 64 DAEIRPWCYTMDPSRWECYCNLTQCPVTSSVLTPTVADVPSTEAPSE 112

RESULT 7
Q02935
ID Q02935 PRELIMINARY; PRT; 290 AA.
AC Q02935;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hepatocyte growth factor, heavy chain precursor.
GN HGF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=91200041; PubMed=1826653;
RA Miyazawa K., Kitamura A., Naka D., Kitamura N.;
RT "An alternatively processed mRNA generated from human hepatocyte
RT growth factor gene";
RL Eur. J. Biochem. 197:15-22(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=93087571; PubMed=1280830;
RA Hartmann G., Naldini L., Weidner K.M., Sachs M., Vigna E.,
RA Comoglio P.M., Birdmeier W.;

```


RT "A functional domain in the heavy chain of scatter factor/hepatocyte growth factor binds the c-Met receptor and induces cell dissociation but not mitogenesis.";
 RT Proc. Natl. Acad. Sci. U.S.A. 89:11574-11578(1992).
 RL FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPIC FACTOR, AND ACTS AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES. IT HAS NO DETECTABLE PROTEASE ACTIVITY.
 CC -!- SIMILARITY: CONTAINS TWO KRINGLE REGIONS.
 CC -!- SIMILARITY: HIGH, TO OTHER HGF; LOWER, TO PLASMINOGEN.
 DR EMBL; X57574; CAA40802.1; -;
 DR EMBL; I02931; AAA52649.1; -;
 DR HSSP; P14210; 1BHT.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan_app.
 DR Pfam; PF00051; Kringle; 2.
 DR Pfam; PF00024; PAN; 1.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00473; PAN.AP; 1.
 DR PROSITE; PS00021; KRINGLE_1; 2.
 DR PROSITE; PS00070; KRINGLE_2; 2.
 DR GlycoProtet; Kringle.
 KW Growth factor; Kringle; Signal.
 FT SIGNAL 1 31 POTENTIAL.
 FT CHAIN 32 290 HEPATOCYTE GROWTH FACTOR, HEAVY CHAIN.
 FT DOMAIN 128 206 KRINGLE 1 (BY SIMILARITY).
 FT DOMAIN 211 288 KRINGLE 2 (BY SIMILARITY).
 FT MOD_RES 32 32 PYRROLIDONE CARBOXYLIC ACID (BY SIMILARITY).
 FT SEQUENCE 290 AA; 33765 MW; C8A18A6FD063200A CRC64;
 Query Match 27.08; Score 481.5; DB 4; Length 290;
 Best local similarity 43.7%; Pred. No. 9e-35;
 Matches 86; Conservative 21; Mismatches 55; Indels 35; Gaps 3;
 QY 5 VDCYHGDSYRGISSTVTGTCQSSSMIPHHQRTPTNPAGLTENYCRNPDGK 64
 DB 125 IRNCILGKGRYKTVSITKSGIKCPQSSMIPHEHSFLPSSYRGKDLQENYCRNP-RGE 183
 QY 65 Q--PWCYTDPCVRYEYCNLTQCSSETSGVLEPTVVPVPSMEAHSEAAPTQTPTVVRQC 122
 DB 184 EGGPWCFTSNPEVRYEYCDIPQCSVE-----C 211
 QY 123 YHNGQSYRGTFSTVTGTCQSSSMTPHRRHQTPTNPNDGLTWNVYCRNPDADTGPWC 182
 DB 212 MTCNGESYRGMDHTSGKICQKWDHQTPHRHKFLPERYPDKGDDNYCRNPDGQPRWC 271
 QY 183 FTTDPSIRWEYCNLTCS 199
 DB 272 YTLDPHTRWEYCAIKTC 288
 RESULT 8
 Q14519 ID Q14519 PRELIMINARY; PRT; 296 AA.
 AC Q14519;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Competitive HGF antagonist.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Chan A.M.-L., Rubin J.S., Bottaro D.P., Hirschfield D.W., Chedid M., Aaronson S.A.;
 RA "Identification of a Competitive HGF Antagonist Encoded by an Alternative Transcript.";
 RT

RL Science 0:0-0(1991).
 CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 DR EMBL; M7727; AAA35980.1; -;
 DR HSSP; P14210; 1BHT.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan_app.
 DR Pfam; PF00051; Kringle; 2.
 DR Pfam; PF00024; PAN; 1.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00473; PAN.AP; 1.
 DR PROSITE; PS00021; KRINGLE_1; 2.
 DR PROSITE; PS00070; KRINGLE_2; 2.
 DR GlycoProtet; Kringle.
 KW Glycoprotein; Kringle.
 SQ SEQUENCE 296 AA; 34546 MW; A45E456B87AE03BE CRC64;
 Query Match 26.7%; Score 477.5; DB 4; Length 296;
 Best local similarity 43.2%; Pred. No. 2.1e-34;
 Matches 86; Conservative 21; Mismatches 57; Indels 35; Gaps 3;
 QY 5 VDCYHGDSYRGISSTVTGTCQSSSMIPHHQRTPTNPAGLTENYCRNPDGK 64
 DB 125 IRNCILGKGRYKTVSITKSGIKCPQSSMIPHEHSFLPSSYRGKDLQENYCRNP-RGE 183
 QY 65 Q--PWCYTDPCVRYEYCNLTQCSSETSGVLEPTVVPVPSMEAHSEAAPTQTPTVVRQC 122
 DB 184 EGGPWCFTSNPEVRYEYCDIPQCSVE-----C 211
 QY 123 YHNGQSYRGTFSTVTGTCQSSSMTPHRRHQTPTNPNDGLTWNVYCRNPDADTGPWC 182
 DB 212 MTCNGESYRGMDHTSGKICQKWDHQTPHRHKFLPERYPDKGDDNYCRNPDGQPRWC 271
 QY 183 FTTDPSIRWEYCNLTCS 201
 DB 272 YTLDPHTRWEYCAIKNRD 290
 RESULT 9
 Q28911 ID Q28911 PRELIMINARY; PRT; 145 AA.
 AC Q28911;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Apolipoprotein[a] (Fragment).
 GN APOLOPROTEIN[A].
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca.
 OC NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95395387; PubMed=7666007;
 RA Ramharack R., Spahr M.A., Hicks G.W., Kieft K.A., Brammer D.W., Minton L.L., Newton R.S.;
 RA "Gemfibrozil significantly lowers cynomolgus monkey plasma lipoprotein[a]-protein and liver apolipoprotein[a] mRNA levels.";
 RT J. Lipid Res. 36:1294-1304(1995).
 CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL; S79621; AAD14312.1; -;
 DR HSSP; P00747; 2PK4.
 DR InterPro; IPR000001; Kringle.
 DR Pfam; PF00051; Kringle; 1.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 KW Glycoprotein; Kringle; Lipoprotein.
 FT NON_TER 145 145

```
SQ SEQUENCE 145 AA; 16097 MW; 0D2DA20A06A5DD87 CRC64;
Query Match 26.2%; Score 467.5; DB 6; Length 145;
Best Local Similarity 64.7%; Pred. No. 7e-34; 25; Indels 11; Gaps 2;
Matches 86; Conservative 11; Mismatches 11;

QY 102 PSMEHSE-----AAPTEQTPVVRQCYHGNGSYRGFTFTVTGTCQSSWMT 151
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 14 PIME-HKEVVLFLKALPEQSHVQDCYHGDGQSYGCTSTTGTGTCQANSMEP 72

QY 152 HRHQRTPENYNDGLTMKNCENADADGPKCETTPDSIRWEYCNLTFCSDTGTGVVAPPT 211
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 73 HOHNRTENYNAGIRNYCRNPDVAAFCYCTMDPNVRWEYCNLTFCSDAEGTAVAPPN 132

QY 212 VIOVPSLGPSPSEQ 224
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 133 VTLVPSLEAPSEQ 145

RESULT 10
Q8TCE2 PRELIMINARY; PRT; 285 AA.
AC Q8TCE2
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Similar to hepatocyte growth factor (hepatopoietin A, scatter factor).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; BC022308; A822308.1; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00024; PAN; 1.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00473; PAN_AP; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
KW Glycoprotein; Kringle.
SQ SEQUENCE 285 AA; 33234 MW; 0A93B073EA86EA61 CRC64;
Query Match 26.1%; Score 466; DB 4; Length 285;
Best Local Similarity 43.1%; Pred. No. 2.1e-33;
Matches 85; Conservative 21; Mismatches 51; Indels 40; Gaps 4;

QY 5 VQDCYHGDGRSGYSGISSTVTGTCQSSWNPHEQRTPEYCNLTFCSDTGTGVVAPPT 64
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 125 IRNCILGKSGYSGIVSITKSGIKQCPNWSMIPHEH----SYRGLQENYCRNP-RGE 178

QY 65 Q-PWCYTPDPCVRWEYCNLTFCSDTGTGVVAPPTVVPVPSMEAHSEAPTEQTPVVRQC 122
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 179 EGGWCTSTNPEYRVEYCDIQCEVE-----C 206

QY 123 VHNGQSYRGFTFTVTGTCQSSWNPHEQRTPEYCNLTFCSDTGTGVVAPPT 182
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 207 MTCNGEYRGLMDHTSGKICQWHDQTPRHFLPERYPDKGDDNYCRNPQGPSPMC 266

QY 183 FTTPDSIRWEYCNLTFC 199
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 267 VTLDPHRWEYCAIKTC 283
```

```
RESULT 11
Q9UIR8 PRELIMINARY; PRT; 105 AA.
AC Q9UIR8
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21181705; PubMed=11285247;
RA Ogorelkova M., Kraft H.G., Ehnholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) Kringles IV
RT types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AFI58656; AAF03677.1; -.
DR EMBL; AFI58655; AAF03677.1; JOINED.
DR HSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 105 105
SQ SEQUENCE 105 AA; 11882 MW; 6ECB6C02CD30EFA2 CRC64;
Query Match 24.9%; Score 445; DB 4; Length 105;
Best Local Similarity 72.5%; Pred. No. 4.7e-32;
Matches 74; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 112 PTEQIPVVRQCYHGNGSYRGFTFTVTGTCQSSWNPHEQRTPEYCNLTFCSDTGTGVVAPPTVI 213
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1 PTEQSPGVQDCYHGDGQSYRGFTFTVTGTCQSSWNPHEQRTPEYCNLTFCSDTGTGVVAPPTVI 60

QY 172 RNPDAETGFCWCTTDPDSIRWEYCNLTFCSDTGTGVVAPPTVI 213
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 61 RNPDAETGFCWCTTDPDSIRWEYCNLTFCSDTGTGVVAPPTVI 102

RESULT 12
Q9N1B8 PRELIMINARY; PRT; 109 AA.
AC Q9N1B8
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hepatocyte growth factor (Fragment).
GN HGF.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Endometrium;
RA MEDLINE=20297031; PubMed=10819791;
RA Chen C., Spencer T.E., Bazer F.W.;
RT "Expression of hepatocyte growth factor and its receptor c-met in the
RT ovine uterus.";
RL Biol. Reprod. 62:1844-1850(2000).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
```

```

DR EMBL; AF213397; AAF25945.1; -.
DR HSSP; P14210; 2HGF.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR Glycoprotein; Kringle.
KW NON_TER 1
FT NON_TER 109
SQ SEQUENCE 109 AA; 12501 MW; 1F88FE5DBC0D4A5D CRC64;

Query Match 14.6%; Score 261.5; DB 6; Length 109;
Best Local Similarity 39.0%; Pred. No. 9.1e-16;
Matches 48; Conservative 15; Mismatches 37; Indels 23; Gaps 2;

Qy 114 EQTPVRCYHNGQSYRGTFSTVTGRTCSQSSMTPPHRHQRTPENYPNDGLTMYCNRN 173
Db 9 ENKDYRNCIIIGKGSYKGTSTKSGIKQCPWNSMIPHEHSLFPSSYRGKDLQENYCNRN 68

Qy 174 PDADT-GPWCFTDPSIRWEYCNLRCSDFTEGVVAPPTVQVPSLGPPSEDCMFCNCK 232
Db 69 PRGEGGPGWCFSTNPEVRYEVCDFPQCSEV-----CMTCNCE 106

Qy 233 GYR 235
Db 107 STR 109

RESULT 13
Q8BS17 PRELIMINARY; PRT; 209 AA.
AC Q8BS17;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Plasminogen.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK040840; BAC30717.1; -.
SQ SEQUENCE 209 AA; 23192 MW; FA5880C13A7FA8E7 CRC64;

Query Match 14.6%; Score 260; DB 11; Length 209;
Best Local Similarity 44.3%; Pred. No. 2.6e-15;
Matches 47; Conservative 15; Mismatches 42; Indels 2; Gaps 2;

Qy 114 EQTPVRCYHNGQSYRGTFSTVTGRTCSQSSMTPPHRHQRTPENYPNDGLTMYCNRN 173
Db 95 EKEVILSECKYIGNGYRGCTMSTKSGVACQKAGTTPHPVNPYSPSTHNEGLGENYCNRN 154

Qy 174 PDADT-GPWCFTDPSIRWEYCNLRCSDFTEGVVAPPTVQVPSL 218
Db 155 PDNDEQGPWCFTDPPDKRYDCNIPCEGRGGS-RKPPSVCLSPGI 199

RESULT 14
Q9BYM0 PRELIMINARY; PRT; 208 AA.
ID Q9BYM0
AC Q9BYM0
DT 01-JUN-2001 (Tremblrel. 17, Created)

```

```

DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hepatocyte growth factor (Fragment).
GN HGF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91369928; PubMed=1832556;
RA Miyazawa K., Kitamura A., Kitamura N.;
RT "Structural organization and the transcription initiation site of the
  human hepatocyte growth factor gene."
RL Biochemistry 30:9170-9176(1991).
CC 1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; M75971; AAG53459.1; -.
DR EMBL; M75971; AAG53459.1; JOINED.
DR EMBL; M75966; AAG53459.1; JOINED.
DR EMBL; M75968; AAG53459.1; JOINED.
DR EMBL; M75969; AAG53459.1; JOINED.
DR HSSP; P14210; 1BHT.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; PAN_app.
DR InterPro; IPR003966; Prothrombin.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00024; PAN; 1.
DR PRINTS; PR00018; KRINGLE
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00473; PAN_AP; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle.
FT NON_TER 208
SQ SEQUENCE 208 AA; 23931 MW; AE9C50DESA86B37B CRC64;

Query Match 14.0%; Score 249.5; DB 4; Length 208;
Best Local Similarity 47.2%; Pred. No. 2.2e-14;
Matches 42; Conservative 14; Mismatches 32; Indels 1; Gaps 1;

Qy 114 EQTPVRCYHNGQSYRGTFSTVTGRTCSQSSMTPPHRHQRTPENYPNDGLTMYCNRN 173
Db 120 ENKDYRNCIIIGKGSYKGTSTKSGIKQCPWNSMIPHEHSLFPSSYRGKDLQENYCNRN 179

Qy 174 PDADT-GPWCFTDPSIRWEYCNLRCSDFTEGVVAPPTVQVPSL 201
Db 180 PRGEGGPGWCFSTNPEVRYEVCDFPQCSE 208

RESULT 15
Q13494 PRELIMINARY; PRT; 210 AA.
ID Q13494
AC Q13494;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE HGF agonist/antagonist.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96278713; PubMed=8662798;
RA Ciocce V., Csaky K.G., Chan A.M.L., Bottaro D.P., Taylor W.G.,
  Jensen R., Aaronson S.A., Rubin J.S.;
RT "Hepatocyte growth factor (HGF)/NK1 is a naturally occurring
  HGF/scatter factor variant with partial agonist/antagonist activity."
RL J. Biol. Chem. 271:13110-13115(1996).
CC 1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

```

```
DR EMBL; U46010; AAC50539.1; -.
DR HSP; P14210; IBHT.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR Pfam; PF00051; kringle; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00473; PAN_AP; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle.
SQ SEQUENCE 210 AA; 24116 MW; 94A6EE9C50DE5A86 CRC64;

Query Match      14.0%; Score 249.5; DB 4; Length 210;
Best Local Similarity 47.2%; Pred. No. 2.3e-14;
Matches 42; Conservative 14; Mismatches 32; Indels 1; Gaps 1;

QY 114 EQTPVVRQCYHGNGSYRGTFSTVTGRTCCSWSSMTGHRHQTPTENYPNDGLTMYCN 173
Db 120 ENKDYIRNCILGKRSYKGTYSITKSGIKQCPWSSMIPHERSFLPSSYRGKDLQENYCN 179

QY 174 PDADT-GPWCFTTDFSIWEYCNLTRGSD 201
Db 180 PRGEGGPWCFTSNPEVAYEVCDIPQCSE 208
```

Search completed: September 29, 2003, 08:01:57
Job time : 40 secs

XX PT Novel angiogenesis inhibitor, LK68 for treating angiogenesis-mediated
PT diseases, e.g. cancer and rheumatoid arthritis, has human
PT apolipoprotein(a) kringle domains IV36, IV37 and V38 amino acid
PT sequence -
XX
XX Claim 1; Page 44-45; 50pp; English.
XX
XX The present sequence is human LK6 protein which contains the amino
CC acid sequence of human apolipoprotein(a) kringle domain IV36
CC (KIV36). The human apolipoprotein(a) kringle domains IV36 (LK6
CC protein), IV37 (LK7 protein) and V38 (LK8 protein) together form the
CC angiogenesis inhibitor, LK68 protein. LK68, LK6, LK7 and LK8 are
CC inhibitors of endothelial cell proliferation, cell migration and
CC normal development of capillaries in the chick embryo chorioallantoic
CC membrane (CAM). LK68 protein, its single kringles or their functional
CC equivalents, are useful for treating angiogenesis-mediated diseases,
CC such as cancer, rheumatoid arthritis, psoriasis or ocular angiogenic
CC disease in animals or humans. LK68 is useful as an anticancer agent and
CC also for inhibiting primary tumour growth.
XX
XX Sequence 91 AA;
SQ Query Match 100.0%; Score 539; DB 22; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.5e-42;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSPVVQCYHGDRGSRYSSTVTGRTCSWSMTPHWHQRTPENYNAGLTENYCRNP 60
DB 1 KSPVVQCYHGDRGSRYSSTVTGRTCSWSMTPHWHQRTPENYNAGLTENYCRNP 60

QY 61 DSGKPCWYCTTDFCVRWEXCNLTQCSSETS 91
DB 61 DSGKPCWYCTTDFCVRWEXCNLTQCSSETS 91

RESULT 2
AAB01909
ID AAB01909 standard; Protein; 90 AA.
XX
XX AAB01909;
XX
XX 18-SEP-2000 (first entry)
XX
XX Human plasminogen kringle 4 (Val354-Val443).
XX
XX Plasminogen; human; kringle domain; endothelial cell proliferation;
KW angiogenesis; antiproliferative; antiarteriosclerotic; cystostatic;
KW antipsoriatic; antiinflammatory; antiulcer; antirheumatic; antiarthritis;
KW antiangiogenic; cancer; tumour; autoimmune disease.
XX
XX Homo sapiens.
XX
XX US6057122-A.
XX
XX 02-MAY-2000.
XX
XX 05-MAY-1997; 97US-0851350.
XX
XX 03-MAY-1996; 96US-0643219.
XX
XX 03-APR-1997; 97US-0832087.
XX
XX (ABBO) ABBOTT LAB.
XX
XX Davidson DJ;
XX
XX WPI; 2000-349573/30.
XX
XX Preparation of Kringle five peptide fragment for treating various
PT disorders such as angiogenic, ocular, skin diseases and cancer,
PT involves mixing mammalian plasminogen and elastase followed by
PT incubation and isolation -
XX

PS Example 17; Page -; 48pp; English.
XX
XX The invention relates to a method of preparing plasminogen kringle 5
CC peptide fragments. The method comprises mixing mammalian plasminogen and
CC elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC endothelial cell proliferation and migration. The peptides are useful
CC for treating angiogenic diseases, primary and metastatic solid tumours
CC and carcinomas of various organs such as breast, genital tract, tumours
CC endocrine glands, skin, tumours of the brain and eyes and solid tumours
CC arising from haematopoietic malignancies such as leukaemias and
CC lymphomas. They are also used for the prophylaxis of various autoimmune
CC diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
CC (e.g., psoriasis), blood vessel diseases (e.g. haemangiomas, Osler-Webber
CC syndrome), diseases caused by excessive or abnormal stimulation of
CC endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
CC which have angiogenesis as a pathologic consequence (e.g., cat scratch
CC disease and ulcers). The peptides are also useful as a birth control
CC agent which inhibits ovulation and establishment of the placenta.
CC Sequences AAB01906-B01919 represent fragments of human plasminogen used
CC in an exemplification of the invention.
CC Note: This sequence is not shown in the specification, but is derived
CC from the full length human plasminogen sequence (AAB01887) shown in
CC figure 1.
XX
XX Sequence 90 AA;
SQ Query Match 77.0%; Score 415; DB 21; Length 90;
Best Local Similarity 79.3%; Pred. No. 4.1e-31;
Matches 69; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 4 VVQDCYHGDGRSYRGISSTVTGRTCSWSMTPHWHQRTPENYNAGLTENYCRNPDSG 63
DB 1 VVQDCYHGDGRSYRGISSTVTGRTCSWSMTPHWHQRTPENYNAGLTENYCRNPDSG 60

QY 64 KQPCWYCTTDFCVRWEXCNLTQCSSETS 90
DB 61 KQPCWYCTTDFCVRWEXCNLTQCSSETS 87

RESULT 3
AAY77720
ID AAY77720 standard; protein; 84 AA.
XX
XX AAY77720;
XX
XX 12-MAY-2000 (first entry)
XX
XX Human plasminogen kringle 4 sequence.
XX
XX Kringle 4; plasminogen; anti-angiogenic; angiogenesis; angiostatin;
KW cytostatic; antiarthritic; antirheumatic; antidiabetic; ophthalmological;
KW immunosuppressant; vasotropic; vulnery; antiarteriosclerotic; human;
KW dermatological; cancer; tumour; birth control; vascularization.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 59..75
XX /note= "specifically claimed fragment (AAY77723)"
XX
XX WO200003726-A1.
XX
XX 27-JAN-2000.
XX
XX 07-JUL-1999; 99WO-US15271.
XX
XX 14-JUL-1998; 98US-0092831.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Ji R, Trail PA;
XX

DR WPI; 2000-171200/15.
XX
XX Novel lysine binding fragments angiotensin used as antiangiogenic
PT agents in the treatment of cancer, diabetic retinopathy, rheumatoid
PT arthritis, psoriasis, atherosclerotic plaque formation, and other
PT angiogenesis diseases -
XX
XX Disclosure; Fig 3; 30pp; English.
XX
XX The invention provides fragments of kringle 1, 2 or 4 of human
CC plasminogen that contain a lysine binding site and have anti-angiogenic
CC activity. The peptides of the invention function as antiangiogenic
CC agents, for the treatment process and diseases involving angiogenesis.
CC Such diseases include cancers such as solid tumours, blood born tumours
CC such as leukemias, tumours metastasis, benign tumours such as
CC hemangiomas, acoustic acuromas, neurofibromas, trachomas and pyogenic
CC granulomas, rheumatoid arthritis, ocular angiogenic diseases such as
CC diabetic retinopathy, retinopathy of prematurity, macular degeneration,
CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasia,
CC rubecosis, Osler-Webber syndrome, myocardial angiogenesis, plaque
CC neovascularization, telangiectasia, hemophilic joints, angiofibroma,
CC and wound granulation. The fragments are also useful in treatment of
CC disease of excessive or abnormal stimulation of endothelial cells. These
CC diseases include intestinal adhesions, atherosclerosis, scleroderma, and
CC hypertrophic scars. The fragments can be used as birth control agents by
CC preventing vascularization required for embryo transplantation. The
CC present sequence represents the kringle 4 sequence of human plasminogen
CC (angiotensin).
XX
XX Sequence 84 AA;
SQ
Query Match 75.1%; Score 405; DB 21; Length 84;
Best Local Similarity 79.8%; Pred. No. 3.2e-30;
Matches 67; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
QY 5 VDCYHGDGRSGISSTVTGTCTQSSMIPHHQRTPEYTNAGLTENYCRNPDSGK 64
DB 1 VDCYHGDGRSGISSTVTGTCTQSSMIPHHQRTPEYTNAGLTENYCRNPDSGK 64
QY 65 QPWCFTTDPVSRWEYCNLTQCSSE 88
DB 61 GPWCFTTDPVSRWEYCNLTQCSSE 84
RESULT 4
AAY72946
ID AAY72946 standard; Protein; 89 AA.
XX
XX AAY72946;
XX
XX 13-JUN-2001 (first entry)
XX
XX Human apolipoprotein(a) kringle domain IV37, LK7 protein.
XX
XX Human; angiogenesis inhibitor; LK7; apolipoprotein(a) kringle domain;
KW angiogenesis-mediated disease; cancer; rheumatoid arthritis; therapy;
KW cytosolic; antithrombotic; antiarthritic; antipsoriasis; psoriasis;
KW ocular angiogenic disease; endothelial cell proliferation; tumour;
KW cell migration.
XX
XX Homo sapiens.
OS
XX WO200119868-A1.
PN
XX 22-MAR-2001.
PD
XX 15-SEP-1999; 99WO-KR00554.
PF
XX 15-SEP-1999; 99WO-KR00554.
PR
XX (MOGA-) MOGAM BIOTECHNOLOGY RES INST.
PA
XX Chang J, Kim JS, Park EJ, Yum J, Chung S;
PI

XX
DR WPI; 2001-244787/25.
DR N-PSDB; AAD03257.
XX
XX Novel angiogenesis inhibitor, LK68 for treating angiogenesis-mediated
PT diseases, e.g. cancer and rheumatoid arthritis, has human
PT apolipoprotein(a) kringle domains IV36, IV37 and V38 amino acid
PT sequence -
XX
XX Claim 2; Page 45-46; 50pp; English.
XX
XX The present sequence is human LK7 protein which contains the amino
CC acid sequence of human apolipoprotein(a) kringle domain IV37
CC (KIV37). The human apolipoprotein(a) kringle domains IV36 (LK6
CC protein), IV37 (LK7 protein) and V38 (LK8 protein) together form the
CC angiogenesis inhibitor, LK68 protein. LK68, LK7 and LK8 are
CC inhibitors of endothelial cell proliferation, cell migration and
CC normal development of capillaries in the chick embryo chorioallantoic
CC membrane (CAM). LK68 protein, its single kringles or their functional
CC equivalents, are useful for treating angiogenesis-mediated diseases,
CC such as cancer, rheumatoid arthritis, psoriasis or ocular angiogenic
CC disease in animals or humans. LK68 is useful as an anticancer agent and
CC also for inhibiting primary tumour growth.
XX
XX Sequence 89 AA;
SQ
Query Match 74.6%; Score 402; DB 22; Length 89;
Best Local Similarity 77.6%; Pred. No. 6.5e-30;
Matches 66; Conservative 8; Mismatches 11; Indels 0; Gaps 0;
QY 5 VDCYHGDGRSGISSTVTGTCTQSSMIPHHQRTPEYTNAGLTENYCRNPDSGK 64
DB 1 VDCYHGDGRSGISSTVTGTCTQSSMIPHHQRTPEYTNAGLTENYCRNPDSGK 64
QY 65 QPWCFTTDPVSRWEYCNLTQCSSE 89
DB 61 GPWCFTTDPVSRWEYCNLTQCSSE 85
RESULT 5
AAP93231
ID AAP93231 standard; peptide; 82 AA.
XX
XX AAP93231;
XX
XX 25-MAR-2003 (updated)
DT
XX 03-APR-1990 (first entry)
DT
XX Plasminogen kringle 4 domain (residues 354-435).
DE
XX Plasminogen; activator; t-PA; fibrin; kringle domain;
KW scu-PA;
KW
XX Homo sapiens.
OS
XX WO8910401-A.
PN
XX 02-NOV-1989.
PD
XX 23-MAR-1989; 89WO-US01255.
PF
XX 22-APR-1988; 88US-0184823.
PR
XX (COLB) COLLABORATIVE RES INC.
PA
XX Mao JI, Abercrombie DM;
PI
XX WPI; 1989-339965/46.
DR
XX N-PSDB; AAN92238.
DR
XX Modified plasminogen activator - having greater fibrin
PT selectivity and circulating half-life.
PT
XX

PS Disclosure; fig 3b; 80pp; English.

XX The sequence is encoded by a synthetic oligonucleotide and is the kringle

CC 4 domain of plasminogen. It is used to construct a modified plasminogen

CC activator (mPA) of greater fibrin selectivity than normal PA. The mPA is

CC used in the treatment of pulmonary embolism, thrombosis, myocardial

CC infarct and strokes.

CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 82 AA;

SQ Query Match 73.7%; Score 397; DB 10; Length 82;

Best Local Similarity 79.3%; Pred. No. 1.7e-29;

Matches 65; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 4 VVQCYHGDSRGYSSTVTGRTCCQSSSMIPHHQRTPENYPNAGLTENYCRNPDG 63

Db 1 VVQCYHNGQSTRGYSSTVTGRTCCQSSSMTPRHQRTPENYPNAGLTENYCRNPD 60

QY 64 KQPCWTTDPCVWEYCNLTQC 85

Db 61 KGPWCFTDPSVWEYCNLKKC 82

RESULT 6

AAW07560

ID AAW07560 standard; protein; 78 AA.

XX

AC AAW07560;

XX

DT 22-JUN-1997 (first entry)

XX

DE Human kringle 4.

XX

KW angiotatin; plasminogen; kringle; angiogenesis; cancer; arthritis;

KW macular degeneration; diabetic retinopathy.

XX

OS Homo sapiens.

XX

PN WO9635774-A2.

XX

PD 14-NOV-1996.

XX

PF 26-APR-1996; 96WO-US05856.

XX

PR 08-MAR-1996; 96US-0612788.

XX

PR 26-APR-1995; 95US-0429743.

XX

PR 22-FEB-1996; 96US-0605598.

XX

PA (CHIL-) CHILDRENS MEDICAL CENT.

XX

PI Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;

XX

PI WPI; 1996-518662/51.

XX

PT Use of angiotatin fragments or aggregates - for inhibiting

PT endothelial cell proliferation and treating angiogenesis-mediated

PT diseases, e.g. cancer, arthritis or diabetic retinopathy

XX

PS Claim 4; Page 116; 203pp; English.

XX

CC The invention relates to new methods and compositions for

CC inhibiting endothelial cell proliferation, using as active component

CC an angiotatin fragment, a combination of angiotatin fragments, or

CC aggregate angiotatin. The fragment is preferably derived from murine,

CC human, Rhesus, porcine or bovine plasminogen and is a kringle 1,

CC kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle

CC 1-4 or kringle 1-4BKLs protein. The aggregate angiotatin has a Mol. Wt.

CC of 45-65 kD and is derived from a plasminogen fragment beginning at

CC approximately amino acid number 98 of murine, human, Rhesus, porcine or

CC bovine plasminogen. The active component can be used for treating

CC angiogenesis-mediated diseases such as cancer, arthritis, macular

CC degeneration and diabetic retinopathy. It can also be used to develop

CC antibodies for use in diagnosis, detection and therapy.

CC The present sequence, human kringle 4, is a specific angiotatin

CC fragment which can be used in the invention.

XX

SQ Sequence 78 AA;

Query Match 71.1%; Score 383; DB 17; Length 78;

Best Local Similarity 79.5%; Pred. No. 3.2e-28;

Matches 62; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 8 CYHGDGRSGYSGISSTVTGRTCCQSSSMIPHHQRTPENYPNAGLTENYCRNPDGKQPW 67

Db 1 CYHGDGSGYSGISSTVTGRTCCQSSSMTPRHQRTPENYPNAGLTENYCRNPDGKQPW 60

QY 68 CYTTDPCVWEYCNLTQC 85

Db 61 CFTDPSVWEYCNLKKC 78

RESULT 7

AA08414

ID AAB08414 standard; Protein; 83 AA.

XX

AC AAB08414;

XX

DT 20-DEC-2000 (first entry)

XX

DE Amino acid sequence of kringle 4 of human plasminogen.

XX

KW Angiogenesis inhibitor; KED; kringle protein; angiogenesis; tPA;

KW tissue plasminogen activator; tumour; atherosclerosis; arthritis;

KW retinopathy; bronchial vascular congestion; inflammatory bowel disease;

KW adult respiratory distress syndrome; Castelman's disease; psoriasis;

KW hepatitis; aneurysm; renal disease; haemangioma.

XX

OS Homo sapiens.

XX

PN WO200049871-A1.

XX

PD 31-AUG-2000.

XX

PF 24-FEB-2000; 2000WO-US04798.

XX

PR 24-FEB-1999; 99US-0121341.

XX

PR 25-FEB-1999; 99US-0121633.

XX

PR 18-NOV-1999; 99US-0166176.

XX

PA (FORD-) FORD HEALTH SYSTEM HENRY.

XX

PI Dou D, Chopp M, Wang L, Mikkelson T;

XX

PI WPI; 2000-572016/53.

XX

PT Use of kringle protein and kringle derived from plasminogen and

PT composition comprising kringle proteins for treating tumor and

PT atherosclerosis, arthritis and retinopathy

XX

PS Disclosure; Fig 6; 163pp; English.

XX

CC The specification describes a human polypeptide which is a potent

CC angiogenesis inhibitor, and is designated KED. KED polypeptide, kringle

CC proteins, or a kringle derived from human tissue plasminogen activator

CC (tPA), protein are used to inhibit angiogenesis. Kringle protein, KED

CC or tPA are useful for treating tumours, as well as atherosclerosis,

CC arthritis, retinopathy and other similar diseases. KED is also useful

CC for the treatment of diseases such as bronchial vascular congestion,

CC inflammatory bowel disease, adult respiratory distress syndrome,

CC Castelman's disease, psoriasis, hepatitis, aneurysm, renal disease

CC and haemangioma. The present sequence represents kringle 4 of human

CC plasminogen, which is used in the course of the invention.

XX

SQ Sequence 83 AA;

Query Match 69.9%; Score 377; DB 21; Length 83;
Best Local Similarity 76.8%; Pred. No. 1.2e-27;
Matches 63; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 5 VDCYHGDSYRGISSTVTGTCOSWSMIPHHQRTPEYFNAGLTENYCRNPDGSK 64
DB 1 VQCYHGDGQSYRGISSTVTGTCOSWSMIPHHQRTPEYFNAGLTENYCRNPDGSK 60

QY 65 QPCYTTDPCVRWEYCNLTQC 86
DB 61 GPTCTDPSVRWEYCNLKCS 82

RESULT 8
AAW07559
ID AAW07559 standard; protein; 78 AA.

AC AAW07559;

XX 22-JUN-1997 (first entry)

DT Murine kringle 4.

DE angiotatin; plasminogen; kringle; angiogenesis; cancer; arthritis;
KW macular degeneration; diabetic retinopathy.

XX Mus musculus.

OS WO9635774-A2.

PN 14-NOV-1996.

XX 26-APR-1996; 96WO-US05856.

PR 08-MAR-1996; 96US-0612788.

PR 26-APR-1995; 95US-0429743.

PR 22-FEB-1996; 96US-0605598.

XX (CHIL-) CHILDRENS MEDICAL CENT.

PA Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;

PI WPI; 1996-518662/51.

XX Use of angiotatin fragments or aggregates - for inhibiting
DR endothelial cell proliferation and treating angiogenesis-mediated
XX diseases, e.g. cancer, arthritis or diabetic retinopathy

PT Claim 4; Page 115; 203pp; English.

PS The invention relates to new methods and compositions for
XX inhibiting endothelial cell proliferation, using as active component
CC an angiotatin fragment, a combination of angiotatin fragments, or
CC aggregate angiotatin. The fragment is preferably derived from murine,
CC human, Rhesus, porcine or bovine plasminogen and is a kringle 1,
CC kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle
CC 1-4 or kringle 1-4BKIS protein. The aggregate angiotatin has a Mol. Wt.
CC of 45-65 kD and is derived from a plasminogen fragment beginning at
CC approximately amino acid number 98 of murine, human, Rhesus, porcine or
CC bovine plasminogen. The active component can be used for treating
CC angiogenesis-mediated diseases such as cancer, arthritis, macular
CC degeneration and diabetic retinopathy. It can also be used to develop
CC antibodies for use in diagnosis, detection and therapy.
CC The present sequence, murine kringle 4, is a specific angiotatin
CC fragment which can be used in the invention, and represents amino
CC acids 377-454 of murine plasminogen.

XX Sequence 78 AA;

Query Match 64.9%; Score 350; DB 17; Length 78;
Best Local Similarity 70.5%; Pred. No. 3.6e-25;
Matches 55; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 8 CYHGDGSRYSRGISSTVTGTCOSWSMIPHHQRTPEYFNAGLTENYCRNPDGSKQFW 67
DB 1 CYSDGQSGRYSRGISSTVTGTCOSWSMIPHHQRTPEYFNAGLTENYCRNPDGSKQFW 60

QY 68 CYTTPDPCVRWEYCNLTQC 85
DB 61 CYTTPDPSVRWEYCNLKRC 78

RESULT 9
ABP02801
ID ABP02801 standard; Protein; 86 AA.

XX ABP02801;

XX 24-JUN-2002 (first entry)

XX Human OREF protein sequence SEQ ID NO:5584.

XX Human; open reading frame; OREF; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis.

XX Homo sapiens.

XX WO200192523-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US10836.

XX 30-MAY-2000; 2000US-206132P.

XX 29-AUG-2000; 2000US-228716P.

XX (CURA-) CURAGEN CORP.

XX Shinkets RA, Leach MD;

XX WPI; 2002-106308/14.

XX N-PSDB; ABN18553.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders

XX Disclosure; SEQ ID 5584; 1037pp; English.

XX The present invention describes substantially purified human proteins
CC (referred to as open reading frame, OREF, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human OREF
CC proteins given in ABP0010 to ABP11500. OREF proteins are useful for
CC treating or preventing a pathology associated with an OREF-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with OREF-associated disorder. OREF polynucleotide
CC sequences can be used in gene therapy. OREF sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. OREF proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: September 29, 2003, 08:18:25 : Search time 29 Seconds
(without alignments)
132.769 Million cell updates/sec

Title: US-10-088-548-4

Perfect score: 539

Sequence: 1 KSPVYDCHGDSYRGIS.....DPCVRYEYCNLTQCSFESG 91

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 224404

Minimum DB seq length: 0

Maximum DB seq length: 91

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2.6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2.6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2.6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2.6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2.6/ptodata/1/iaa/PCTUS_COMB.pep.*

6: /cgn2.6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	405	75.1	84	US-09-348-953-2	Sequence 2, Appli
2	383	71.1	78	US-08-612-788-23	Sequence 23, Appl
3	383	71.1	78	US-08-612-788-23	Sequence 5, Appli
4	383	71.1	78	US-09-066-028-23	Sequence 23, Appl
5	383	71.1	78	US-09-335-325-23	Sequence 23, Appl
6	350	64.9	78	US-08-612-788-22	Sequence 22, Appl
7	350	64.9	78	US-09-066-028-22	Sequence 22, Appl
8	350	64.9	78	US-09-335-325-22	Sequence 22, Appl
9	285	52.9	78	US-08-612-788-17	Sequence 17, Appl
10	285	52.9	78	US-09-066-028-17	Sequence 17, Appl
11	285	52.9	78	US-09-335-325-17	Sequence 17, Appl
12	281	52.1	78	US-08-612-788-18	Sequence 18, Appl
13	281	52.1	78	US-08-612-788-19	Sequence 19, Appl
14	281	52.1	78	US-09-066-028-18	Sequence 18, Appl
15	281	52.1	78	US-09-066-028-19	Sequence 19, Appl
16	281	52.1	78	US-09-335-325-18	Sequence 19, Appl
17	281	52.1	78	US-09-335-325-19	Sequence 18, Appl
18	278	51.6	78	US-08-612-788-20	Sequence 20, Appl
19	278	51.6	78	US-09-066-028-20	Sequence 20, Appl
20	278	51.6	78	US-09-335-325-20	Sequence 20, Appl
21	277	51.4	78	US-08-612-788-4	Sequence 4, Appli
22	273	50.6	78	US-08-612-788-21	Sequence 21, Appl
23	273	50.6	78	US-09-066-028-21	Sequence 21, Appl
24	273	50.6	78	US-09-335-325-21	Sequence 21, Appl
25	251	46.6	90	US-08-643-219-4	Sequence 4, Appli
26	251	46.6	90	US-09-131-995-4	Sequence 4, Appli
27	251	46.6	90	US-08-832-087B-4	Sequence 4, Appli

```

28 251 46.6 90 3 US-09-132-154-4 Sequence 4, Appli
29 250.5 46.5 79 2 US-08-612-788-9 Sequence 9, Appli
30 250.5 46.5 79 3 US-09-066-028-9 Sequence 9, Appli
31 250.5 46.5 79 4 US-09-335-325-9 Sequence 9, Appli
32 249.5 46.3 79 2 US-08-612-788-11 Sequence 11, Appli
33 249.5 46.3 79 3 US-09-066-028-11 Sequence 11, Appli
34 249.5 46.3 79 4 US-09-335-325-11 Sequence 11, Appli
35 249.5 46.3 82 4 US-09-348-953-1 Sequence 1, Appli
36 248 46.0 78 2 US-08-612-788-16 Sequence 16, Appli
37 248 46.0 78 3 US-09-066-028-16 Sequence 16, Appli
38 248 46.0 78 4 US-09-335-325-16 Sequence 16, Appli
39 247.5 45.9 79 2 US-08-612-788-8 Sequence 8, Appli
40 247.5 45.9 79 3 US-08-612-788-10 Sequence 10, Appli
41 247.5 45.9 79 2 US-08-763-528A-2 Sequence 2, Appli
42 247.5 45.9 79 3 US-09-066-028-8 Sequence 8, Appli
43 247.5 45.9 79 3 US-09-066-028-10 Sequence 10, Appli
44 247.5 45.9 79 4 US-09-335-325-8 Sequence 8, Appli
45 247.5 45.9 79 4 US-09-335-325-10 Sequence 10, Appli

```

ALIGNMENTS

RESULT 1

US-09-348-953-2

; Sequence 2, Application US/09348953

; Patent No. 6538103

; GENERAL INFORMATION:

; APPLICANT: JI, Richard Wei-Dong

; APPLICANT: Trail, Pamela A.

; TITLE OF INVENTION: LYSINE BINDING FRAGMENTS OF ANGIOSTATIN

; FILE REFERENCE: DS1Sequences

; CURRENT APPLICATION NUMBER: US/09/348,953

; CURRENT FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: 60/092,831

; PRIOR FILING DATE: 1998-07-14

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 84

; TYPE: PRT

; ORGANISM: human

US-09-348-953-2

```

Query Match          75.18; Score 405; DB 4; Length 84;
Best Local Similarity 79.8%; Pred. No. 1.3e-36;
Matches 67; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

```

QY 5 VDCYHGDGRSYRGISSTVTGTCSQSSMIPHWHQRTPEYFNAGLTENYCRNPDGK 64

Db 1 VDCYHGDGQSYRGTSSTTTGKKQSSWMTPHRHQKTPENYFNAGLTMYCRNPDADK 60

QY 65 QPWCYTTPDCVRYEYCNLTQCSF 88

Db 61 GPWCFTTDPVRYEYCNLKCSGT 84

RESULT 2

US-08-612-788-23

; Sequence 23, Application US/08612788

; Patent No. 5837682

; GENERAL INFORMATION:

; APPLICANT: Folkman, M. Judah

; APPLICANT: O'Reilly, Micheal

; APPLICANT: Cao, Yihai

; APPLICANT: Sim, B. Kim Lee

; TITLE OF INVENTION: Angiostatin Fragments and Method of Use

; NUMBER OF SEQUENCES: 45

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Jones & Askew

; STREET: 191 Peachtree Street, 37th Floor

; CITY: Atlanta

; STATE: Georgia

```
/
/ COUNTRY: U.S.
/ ZIP: 30303-1769
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE:
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 23:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 78 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ IMMEDIATE SOURCE:
/ CLONE: K4
/
/ US-08-612-788-23
/
/ Query Match 71.1%; Score 383; DB 2; Length 78;
/ Best Local Similarity 79.5%; Pred. No. 2.9e-34;
/ Matches 62; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
/
/ QY 8 CYHGDGRSYRGISSTVTGRTCCQSSSMIPWHQRTPEPNYPNAGLTENYCRNPDSGKQPW 67
/ Db 1 CYHGDGSGYRGTSSTTTGKRCQSSSMTPHRHQKTPENYPNAGLTMYCRNPDAKGPW 60
/
/ QY 68 CYTTDPCVRWEYCNLFQC 85
/ Db 61 CFTTDPSPVRWEYCNLKKC 78
/
/ RESULT 3
/ US-08-763-528A-5
/ Sequence 5, Application US/08763528A
/ Patent No. 5854221
/ GENERAL INFORMATION:
/ APPLICANT: Folkmann, M. Judah
/ APPLICANT: Cao, Yihai
/ TITLE OF INVENTION: Endothelial Cell Proliferation Inhibitor
/ TITLE OF INVENTION: and Method of Use
/ NUMBER OF SEQUENCES: 6
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew, LLP
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: US
/ ZIP: 30303
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/763,528A
/ FILING DATE: 12-DEC-1996
/ CLASSIFICATION: 530
/
```

```
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05940-0251
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 5:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 78 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: internal
/ FEATURE:
/ NAME/KEY: Protein
/ LOCATION: 1..78
/ OTHER INFORMATION: /note= "Kringle 4 - Figure 3"
/
/ US-08-763-528A-5
/
/ Query Match 71.1%; Score 383; DB 2; Length 78;
/ Best Local Similarity 79.5%; Pred. No. 2.9e-34;
/ Matches 62; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
/
/ QY 8 CYHGDGRSYRGISSTVTGRTCCQSSSMIPWHQRTPEPNYPNAGLTENYCRNPDSGKQPW 67
/ Db 1 CYHGDGSGYRGTSSTTTGKRCQSSSMTPHRHQKTPENYPNAGLTMYCRNPDAKGPW 60
/
/ QY 68 CYTTDPCVRWEYCNLFQC 85
/ Db 61 CFTTDPSPVRWEYCNLKKC 78
/
/ RESULT 4
/ US-09-066-028-23
/ Sequence 23, Application US/09066028
/ Patent No. 6024888
/ GENERAL INFORMATION:
/ APPLICANT: Folkmann, M. Judah
/ APPLICANT: O'Reilly, Micheal
/ APPLICANT: Cao, Yihai
/ APPLICANT: Sim, B. Kim Lee
/ TITLE OF INVENTION: Angiostatin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S.
/ ZIP: 30303-1769
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/066,028
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/612,788
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/
```

```

/ INFORMATION FOR SEQ ID NO: 23:
/
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 78 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHEetical: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ IMMEDIATE SOURCE:
/ CLONE: K4
/ US-09-056-028-23

```

```

Query Match      71.1%; Score 383; DB 3; Length 78;
Best local Similarity 79.5%; Pred. No. 2.9e-34;
Matches 62; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY      8 CYHGGRSYRGISSTVTGRCQSWSSMTPHHQRTPEPNAGLTGTCRNPDSGKQFW 67
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 CYHGDQSYRGTSSTTTGKKCQSWSSMTPHHQRTPEPNAGLTGTCRNPDSGKFW 60
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      68 CYTTDPCVREWCNLTQC 85
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61 CFTTDPSEVREWCNKKC 78
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

;
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: K4
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-335-325-23

Query Match          71.1%; Score 383; DB 4; Length 78;
Best Local Similarity 79.5%; Pred. No. 2.9e-34;
Matches 62; Conservative 6; Mismatches 10; Indels 0; Gaps 0

QY      8 CYHGGRSGVEGSSVTYGRFCQSSKSWIHWORTPENYNAGLTENYCNRPDPSGQKW 67
Db      1 CYHGGOSYRGVSTTTGKKCCQSSWMTFRRHQKTPENYNAGLTMYCNRPDADKGPW 60

QY      68 CYTTDPCVRWEYCNLTQC 85
Db      61 CFTTDPVRWEYCNLKKC 78

RESULT 6
US-08-612-788-22
; sequence 22, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judeah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yifan
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatenIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K4
US-08-612-788-22

```

```

QY      68  CYTTDPCVRWEYCNLTQC 85
      ||||| ||||| ||||| :
Db      61  CYTTDPSVRWEYCNLKRC 78

RESULT 8
US-09-335-325-22
; Sequence 22, Application US/09335325
; Patent No. 6521439
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; O'Reilly, Micheal
; Cao, Yihai
; Sim, B. Kim Lee
; TITLE OF INVENTION: Angiotatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K4
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-335-325-22

      Query Match          64.9%; Score 350; DB 4; Length 78;
      Best Local Similarity 70.5%; Pred. No. 1e-30;
      Matches 55; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY      8  CYHGDGRYSGISSTVTGRTQSWSSMIPHWHQRTPENYPNAGLTENYCRNPDGSKQFW 67
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1  CYQSDQSGYSGYSTTITGKKQSWAAMPHRHSXTPENFDPAGLENNYCRNPDGDKGFW 60

QY      68  CYTTDPCVRWEYCNLTQC 85
      ||||| ||||| ||||| :
Db      61  CYTTDPSVRWEYCNLKRC 78

```



```
RESULT 9
US-08-612-788-17
; Sequence 17, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K3
;
US-08-612-788-17
; Query Match 52.9%; Score 285; DB 2; Length 78;
; Best Local Similarity 59.0%; Pred. No. 1e-23;
; Matches 46; Conservative 6; Mismatches 26; Indels 0; Gaps 0;
QY 8 CYHGDGRYRGISSTVTGTCQSWSMIPHWQRTPEYFNAGLTENYCRNPDGKQPW 67
Db 1 CLKRGENYRGTVSVTSGTCQWSEQTQPHRHRTPEYFNAGLTENYCRNPDGKQPW 60
QY 68 CYTTDPQVRYEYCNLTQC 85
Db 61 CYTTDSQLRWYCEIPSC 78
;
; RESULT 10
US-09-066-028-17
; Sequence 17, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
```

```
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K3
;
US-09-066-028-17
; Query Match 52.9%; Score 285; DB 3; Length 78;
; Best Local Similarity 59.0%; Pred. No. 1e-23;
; Matches 46; Conservative 6; Mismatches 26; Indels 0; Gaps 0;
QY 8 CYHGDGRYRGISSTVTGTCQSWSMIPHWQRTPEYFNAGLTENYCRNPDGKQPW 67
Db 1 CLKRGENYRGTVSVTSGTCQWSEQTQPHRHRTPEYFNAGLTENYCRNPDGKQPW 60
QY 68 CYTTDPQVRYEYCNLTQC 85
Db 61 CYTTDSQLRWYCEIPSC 78
;
; RESULT 11
US-09-335-325-17
; Sequence 17, Application US/09335325
; Patent No. 6521439
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
```



```
;
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 78 amino acids
;   TYPE: amino acid
;   STRANDEDNESS:
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
;   ORGANISM: Rhesus monkey
;   IMMEDIATE SOURCE:
;   CLONE: K3
; US-08-612-788-19
;
; Query Match 52.1%; Score 281; DB 2; Length 78;
; Best Local Similarity 57.7%; Pred. No. 2.7e-23;
; Matches 45; Conservative 7; Mismatches 26; Indels 0; Gaps 0;
;
; QY 8 CYHGDGRSYRGISSTVTGTCQSSMIPWHQRTPEPNYPNAGLTENYCRNPDGKQPW 67
; DB 1 CLKGTGENYRGDVAVTVSGHTCHGWSAQTPHTHRTPEPNFCKNLNDENYCRNPDGKAPW 60
;
; QY 68 CYTTDPCVRWEYCNLTQC 85
; DB 61 CHTINSQVRWEYCKIPSC 78
;
; RESULT 14
; US-09-066-028-18
; Sequence 18, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
;   APPLICANT: Folkman, M. Judah
;   APPLICANT: O'Reilly, Micheal
;   APPLICANT: Cao, Yihai
;   APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Jones & Askew
;   STREET: 191 Peachtree Street, 37th Floor
;   CITY: Atlanta
;   STATE: Georgia
;   COUNTRY: U.S.
;   ZIP: 30303-1769
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/066,028
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/612,788
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Warren, William L.
;     REGISTRATION NUMBER: 36,714
;     REFERENCE/DOCKET NUMBER: 05213-0126
;     TELEPHONE: 404-818-3700
;     TELEFAX: 404-818-3799
;     INFORMATION FOR SEQ ID NO: 18:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH: 78 amino acids
;         TYPE: amino acid
;         STRANDEDNESS:
;         TOPOLOGY: linear
;
; QY 8 CYHGDGRSYRGISSTVTGTCQSSMIPWHQRTPEPNYPNAGLTENYCRNPDGKQPW 67
; DB 1 CLKGTGENYRGDVAVTVSGHTCHGWSAQTPHTHRTPEPNFCKNLNDENYCRNPDGKAPW 60
;
; QY 68 CYTTDPCVRWEYCNLTQC 85
; DB 61 CHTINSQVRWEYCKIPSC 78
;
; RESULT 15
; US-09-066-028-19
; Sequence 19, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
;   APPLICANT: Folkman, M. Judah
;   APPLICANT: O'Reilly, Micheal
;   APPLICANT: Cao, Yihai
;   APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Jones & Askew
;   STREET: 191 Peachtree Street, 37th Floor
;   CITY: Atlanta
;   STATE: Georgia
;   COUNTRY: U.S.
;   ZIP: 30303-1769
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/066,028
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/612,788
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Warren, William L.
;     REGISTRATION NUMBER: 36,714
;     REFERENCE/DOCKET NUMBER: 05213-0126
;     TELEPHONE: 404-818-3700
;     TELEFAX: 404-818-3799
;     INFORMATION FOR SEQ ID NO: 19:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH: 78 amino acids
;         TYPE: amino acid
;         STRANDEDNESS:
;         TOPOLOGY: linear
;         MOLECULE TYPE: protein
;         HYPOTHETICAL: NO
;         ANTI-SENSE: NO
;         FRAGMENT TYPE: N-terminal
;         ORIGINAL SOURCE:
;           ORGANISM: Rhesus monkey
;           IMMEDIATE SOURCE:
```

CLONE: K3
US-09-066-028-19

Query Match 52.1%; Score 281; DB 3; Length 78;
Best Local Similarity 57.7%; Pred. No. 2.7e-23;
Matches 45; Conservative 7; Mismatches 26; Indels 0; Gaps 0;
QY 8 CYHGDGRSYRGISSTTVGRTCCQSSSMIPHWHTPENYPNAGLTENYCRNPDGKQFW 67
Db 1 CLKGTGENYRGDVAVTVSGHTCHGWSAQTPHTHTNTPENFPCKNLDENYCRNPDGKAPW 60
QY 68 CYTDPQVRWEYCNLTQC 85
Db 61 CYTNSQVRWEYCKIPSC 78

Search completed: September 29, 2003, 08:23:21
Job time : 31 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 Compugen Ltd.
 OM protein - protein search, using sw model
 Run on: September 29, 2003, 08:22:08 : Search time 63 Seconds
 (without alignments)
 218.555 Million cell updates/sec

Title: US-10-088-548-4
 Perfect score: 539
 Sequence: 1 KSPVQDCYHGDSRYGIS.....DPCRWECNLTQCSHESG 91

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 203300

Minimum DB seq length: 0

Maximum DB seq length: 91

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*
 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	383	71.1	78	9	US-09-753-064-5
2	383	71.1	78	9	US-09-761-120-23
3	383	71.1	78	10	US-09-335-325-23
4	383	71.1	78	12	US-10-267-137-9
5	383	71.1	78	15	US-10-131-241-23
6	350	64.9	78	9	US-09-761-120-22
7	350	64.9	78	10	US-09-335-325-22
8	350	64.9	78	15	US-10-131-241-22
9	285	52.9	78	9	US-09-761-120-17
10	285	52.9	78	10	US-09-335-325-17
11	285	52.9	78	15	US-10-131-241-17
12	281	52.1	78	9	US-09-761-120-19
13	281	52.1	78	10	US-09-335-325-18
14	281	52.1	78	10	US-09-335-325-19
15	281	52.1	78	15	US-10-131-241-18

16	281	52.1	78	15	US-10-131-241-19
17	278	51.6	78	9	US-09-761-120-20
18	278	51.6	78	10	US-09-335-325-20
19	278	51.6	78	15	US-10-131-241-20
20	277	51.4	78	9	US-09-753-064-4
21	277	51.4	78	9	US-09-761-120-18
22	273	50.6	78	9	US-09-761-120-21
23	273	50.6	78	10	US-09-335-325-21
24	273	50.6	78	15	US-10-131-241-21
25	271	50.3	78	12	US-10-267-137-12
26	250.5	46.5	79	9	US-09-761-120-9
27	250.5	46.5	79	10	US-09-335-325-9
28	250.5	46.5	79	15	US-10-131-241-9
29	249.5	46.3	79	9	US-09-761-120-11
30	249.5	46.3	79	10	US-09-335-325-11
31	249.5	46.3	79	15	US-10-131-241-11
32	248	46.0	78	9	US-09-761-120-16
33	248	46.0	78	10	US-09-335-325-16
34	248	46.0	78	15	US-10-131-241-16
35	247.5	45.9	79	9	US-09-753-064-2
36	247.5	45.9	79	9	US-09-761-120-8
37	247.5	45.9	79	9	US-09-761-120-10
38	247.5	45.9	79	10	US-09-335-325-8
39	247.5	45.9	79	10	US-09-335-325-10
40	247.5	45.9	79	12	US-10-267-137-7
41	247.5	45.9	79	15	US-10-131-241-8
42	247.5	45.9	79	15	US-10-131-241-10
43	245.5	45.5	79	10	US-09-335-325-7
44	245.5	45.5	79	15	US-10-131-241-7
45	244	45.3	78	9	US-09-761-120-15

ALIGNMENTS

RESULT 1

US-09-753-064-5
 : Sequence 5, Application US/09753064
 : Patent No. US20010016644A1
 : GENERAL INFORMATION:
 : APPLICANT: Cao, Yihai
 : Folkman, M. Judah
 : TITLE OF INVENTION: Endothelial Cell Proliferation Inhibitor
 : and Method of Use
 : NUMBER OF SEQUENCES: 6
 : CORRESPONDENCE ADDRESSES:
 : ADDRESSEE: Jones & Askew, LLP
 : STREET: 191 Peachtree Street, 37th Floor
 : CITY: Atlanta
 : STATE: Georgia
 : COUNTRY: US
 : ZIP: 30303
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/753,064
 : FILING DATE: 29-Dec-2000
 : CLASSIFICATION: <Unknown>
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/763,528
 : FILING DATE: 12-DEC-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Warren, William L.
 : REGISTRATION NUMBER: 36,714
 : REFERENCE/DOCKET NUMBER: 05940-0251
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 404-818-3700
 : TELEFAX: 404-818-3799
 : INFORMATION FOR SEQ ID NO: 5:
 : SEQUENCE CHARACTERISTICS:

```

? O'Reilly, Micheal
? Cao, Yinhai
? Sim, B. Kim Lee
? TITLE OF INVENTION: Angiotatin Fragments and Method of Use
? NUMBER OF SEQUENCES: 45
? CORRESPONDENCE ADDRESS:
? ADDRESSES: Jones & Askew
? STREET: 191 Peachtree Street, 37th Floor
? CITY: Atlanta
? STATE: Georgia
? COUNTRY: U.S.
? ZIP: 30303-1769
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/335,325
? FILING DATE: 17-Jun-1999
? CLASSIFICATION: <Unknown>
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US/08/612,788
? FILING DATE: <Unknown>
? ATTORNEY/AGENT INFORMATION:
? NAME: Warren, William L.
? REGISTRATION NUMBER: 36,714
? REFERENCE/DOCKET NUMBER: 05213-0126
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 404-818-3700
? TELEFAX: 404-818-3799
? INFORMATION FOR SEQ ID NO: 23:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 78 amino acids
? TYPE: amino acid
? STRANDEDNESS: <Unknown>
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? FRAGMENT TYPE: N-terminal
? ORIGINAL SOURCE:
? ORGANISM: Homo sapiens
? IMMEDIATE SOURCE:
? CLONE: K4
? SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-335-325-23
Query Match 71.1%; Score 383; DB 10; Length 78;
Best Local Similarity 79.5%; Pred. No. 3.8e-34;
Matches 62; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
QY 8 CYHGDGRSYRGISSTVTGRTCCWSMIPHHPHORTPENYNAGLTENYCRNPDGSKQPW 67
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CYHGDSGVSGYSTTTTGKKGCSWSMPPHRHXPENTPENYNAGLTNYCRNPDAKGWP 60
QY 68 CYTDCPVREYCNLTQC 85
| | | | | | | | | | | :
Db 61 CFTDPSVRWEYCNLKKC 78
RESULT 4
US-10-267-137-9
? Sequence 9, Application US/10267137
? Publication No. US20030148950A1
? GENERAL INFORMATION:
? APPLICANT: Xin, Li
? APPLICANT: Li, Zai-Ping
? APPLICANT: Gan, Ren-bao
? APPLICANT: Zhou, Qing-wei
? APPLICANT: Xu, Ren
? TITLE OF INVENTION: KRINGLE DOMAIN 1 OF HUMAN HEPATOCYTE
? GROWTH FACTOR AND USE THEREFOR
```

```
; FILE REFERENCE: 524282000100
; CURRENT APPLICATION NUMBER: US/10/267,137
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 60/328,329
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence alignment of kringle domains of
; OTHER INFORMATION: plasminogen and HGF
US-10-267-137-9

Query Match          71.1%; Score 383; DB 12; Length 78;
Best Local Similarity 79.5%; Pred. No. 3.8e-34;
Matches 62; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 8 CYHGDGRSYRGISSTVTGRTCSWSSMIPHWHQRTPENYPNAGLTENYCRNPDGKQPW 67
   |||||:|||| |||| ||: ||||| || |||||:|||| |||||: ||
Db 1 CYHGDGQSYRGTSSTTTGKCCQSWSSMTPHRRHQXTPENYPNAGLTWNYCRNPDADKGPW 60

QY 68 CYTTDPCVWEYCNLTQC 85
   |:|||| |||||:||||:|
Db 61 CFTTDPVSRWEYCNLKRC 78

RESULT 5
US-10-131-241-23
; Sequence 23, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; TITLE OF INVENTION: Compositions and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-23

Query Match          71.1%; Score 383; DB 15; Length 78;
Best Local Similarity 79.5%; Pred. No. 3.8e-34;
Matches 62; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 8 CYHGDGRSYRGISSTVTGRTCSWSSMIPHWHQRTPENYPNAGLTENYCRNPDGKQPW 67
   |||||:|||| |||| ||: ||||| || |||||:|||| |||||: ||
Db 1 CYHGDGQSYRGTSSTTTGKCCQSWSSMTPHRRHQXTPENYPNAGLTWNYCRNPDADKGPW 60

QY 68 CYTTDPCVWEYCNLTQC 85
   |:|||| |||||:||||:|
Db 61 CFTTDPVSRWEYCNLKRC 78

RESULT 6
US-09-761-120-22
; Sequence 22, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmi
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Kringle 4
US-09-761-120-22

Query Match          64.9%; Score 350; DB 9; Length 78;
Best Local Similarity 70.5%; Pred. No. 1.4e-30;
Matches 55; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 8 CYHGDGRSYRGISSTVTGRTCSWSSMIPHWHQRTPENYPNAGLTENYCRNPDGKQPW 67
   |||||:|||| |||| ||: ||||| || |||||:|||| |||||: ||
Db 1 CYHGDGQSYRGTSSTTTGKCCQSWAMFPHRSKTFPENFPDAGLEMMYCRNPDGKGPW 60

QY 68 CYTTDPCVWEYCNLTQC 85
   ||||| |||||:||||:|
Db 61 CFTTDPVSRWEYCNLKRC 78

RESULT 7
US-09-335-325-22
; Sequence 22, Application US/09335325
; Patent No. US20020164717A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K4
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-335-325-22

Query Match          54.9%; Score 350; DB 10; Length 78;
Best Local Similarity 70.5%; Pred. No. 1.4e-30;
Matches 55; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 8 CYHGDGRSYRGISSTVTGRTCSWSSMIPHWHQRTPEPNYAGLTENYCRNPDGKQPW 67
DB 1 CYSDGOSYRGTSSTITGKQSWAAMPFRRHSKTPENFPDAGLEMYCRNPDGKGPW 60

QY 68 CYTTDPCVRWEYCNLTQC 85
DB 61 CYTTDPSVRWEYCNLKRC 78

RESULT 8
US-10-131-241-22
; Sequence 22, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Murinae sp.
US-10-131-241-22

Query Match          54.9%; Score 350; DB 15; Length 78;
Best Local Similarity 70.5%; Pred. No. 1.4e-30;
Matches 55; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 8 CYHGDGRSYRGISSTVTGRTCSWSSMIPHWHQRTPEPNYAGLTENYCRNPDGKQPW 67
DB 1 CYSDGOSYRGTSSTITGKQSWAAMPFRRHSKTPENFPDAGLEMYCRNPDGKGPW 60

QY 68 CYTTDPCVRWEYCNLTQC 85
DB 61 CYTTDPSVRWEYCNLKRC 78

RESULT 9
US-09-761-120-17
; Sequence 17, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
```

```
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmi
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Kringle 3
US-09-761-120-17

Query Match          52.9%; Score 285; DB 9; Length 78;
Best Local Similarity 59.0%; Pred. No. 1.5e-23;
Matches 46; Conservative 6; Mismatches 26; Indels 0; Gaps 0;

QY 8 CYHGDGRSYRGISSTVTGRTCSWSSMIPHWHQRTPEPNYAGLTENYCRNPDGKQPW 67
DB 1 CLKRGENTRGTVSVTVSGKTCQWSEQFPHRNTPEPFCKNLENYCRNPDGETAPW 60

QY 68 CYTTDPCVRWEYCNLTQC 85
DB 61 CYTTDQQLRWEYCEIPSC 78

RESULT 10
US-09-335-325-17
; Sequence 17, Application US/09335325
; Patent No. US20020164717A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; O'Reilly, Micheal
; Cao, Fihai
; Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
```



```

; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K3
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-335-325-17

Query Match          52.9%; Score 285; DB 10; Length 78;
Best Local Similarity 59.0%; Pred. No. 1.5e-23;
Matches 46; Conservative 6; Mismatches 26; Indels 0; Gaps 0;

QY      8 CYHGDGRSYRGISSTVTGRTCSQSSMIPHWHQRTPENYPNAGLTENCRNPDSKGQPW 67
Db      1 CLKGRGENYRGVTSVTSKTKCQWSQTTHRHNRTPENFPCKNLLENYCNRPDGTAPW 60
        | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      68 CYYTDPVCVRWEYCNIQTQC 85
Db      61 CYYTDSLRLWEYCEIPSC 78
        ||||| : ||||| : |

RESULT 11
US-10-131-241-17
; Sequence 17, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131.241
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Murine
US-10-131-241-17

Query Match          52.9%; Score 285; DB 15; Length 78;
Best Local Similarity 59.0%; Pred. No. 1.5e-23;
Matches 46; Conservative 6; Mismatches 26; Indels 0; Gaps 0;

QY      8 CYHGDGRSYRGISSTVTGRTCSQSSMIPHWHQRTPENYPNAGLTENCRNPDSKGQPW 67
Db      1 CLKGRGENYRGVTSVTSKTKCQWSQTTHRHNRTPENFPCKNLLENYCNRPDGTAPW 60
        | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      68 CYYTDPVCVRWEYCNIQTQC 85
Db      61 CYYTDSLRLWEYCEIPSC 78
        ||||| : ||||| : |

RESULT 12
US-09-761-120-19
; Sequence 19, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:12:27 ; Search time 40 Seconds
(without alignments)
218.784 Million cell updates/sec

Title: US-10-088-548-4
Perfect score: 539
Sequence: 1 KSPVQDCYHGDRSYRGIS.....DFCVRWEYCNLTQCSFESG 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 32085

Minimum DB seq length: 0
Maximum DB seq length: 91

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	329	61.0	89	2 A60140	plasmin (EC 3.4.21
2	66	12.2	86	2 S20651	Ig heavy chain V r
3	57.5	10.7	77	2 T14395	proteinase inhibit
4	55.5	10.3	86	2 A37910	muscarinic toxin 2
5	55.5	10.3	88	2 S16161	BT14 protein - ba
6	55	10.2	86	2 S26459	Ig kappa chain V r
7	55	10.2	86	2 S29587	Ig kappa chain V r
8	54	10.0	57	1 T2N7BE	short toxin CM-1b
9	52.5	9.7	67	2 S4216	Ig kappa chain - m
10	51.5	9.6	55	2 D9636	unknown protein, 8
11	51	9.5	88	2 G69136	ribosomal protein
12	50	9.3	72	2 D8557	hypothetical prote
13	50	9.3	76	1 IHER2	high potential iro
14	49.5	9.2	82	2 JC5892	cobrotoxin b - Chi
15	49.5	9.2	84	2 P6C142	replication-associ
16	49	9.1	72	1 Q3ECFS	hypothetical 8k pr
17	49	9.1	81	1 IPDK	insulin precursor
18	48.5	9.0	72	1 N2EP1V	long neurotoxin 1
19	48.5	9.0	77	2 S30578	proteinase inhibitor
20	48.5	9.0	84	2 E86209	protein F2G5.16 [
21	48.5	9.0	86	2 S78488	Ig kappa chain V r
22	48.5	9.0	89	2 P10129	class II histocomp
23	48.5	9.0	89	2 I68555	MHC class II histo
24	48	8.9	87	2 B60549	kappa-3 neurotoxin
25	48	8.9	89	2 S38704	MHC class II histo
26	47.5	8.8	55	2 S34330	sox protein - less
27	47.5	8.8	62	1 N1N3M	short neurotoxin 3
28	47.5	8.8	77	2 B84433	proteinase inhibit
29	47.5	8.8	86	2 S34086	Ig kappa chain V r

30	47.5	8.8	87	2 S34084	Ig kappa chain V r
31	47.5	8.8	87	2 S34083	Ig kappa chain V r
32	47.5	8.8	88	2 S21528	Ig kappa chain V r
33	47.5	8.8	88	2 S21522	Ig kappa chain V r
34	47	8.7	55	2 E58892	H+-transporting tw
35	47	8.7	62	2 G25866	short neurotoxin b
36	47	8.7	66	2 P00342	spike glycoprotein
37	47	8.7	66	2 P00340	spike glycoprotein
38	47	8.7	70	2 T32946	hypothetical prote
39	47	8.7	73	2 A1043	transcription regu
40	47	8.7	87	2 S16842	Ig kappa chain V r
41	47	8.7	87	2 S16841	Ig kappa chain V r
42	46.5	8.6	77	2 C84433	proteinase inhibit
43	46.5	8.6	81	2 A11441	intestinal trefoil
44	46	8.5	61	1 N1R12	short neurotoxin 2
45	46	8.5	76	2 T34687	hypothetical prote

ALIGNMENTS

RESULT 1

A60140
plasmin (EC 3.4.21.7) precursor - chicken (fragment)
N;alternate names: plasminogen
C;Species: Gallus gallus (chicken)
C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 16-Jul-1999
C;Accession: A60140

R;Genes, M.; Patthy, L.
Biochim. Biophys. Acta 832, 326-330, 1985
A;Title: The kringle 4 domain of chicken plasminogen.
A;Reference number: A60140; PMID:86077796; PMID:4074753
A;Accession: A60140
A;Molecule type: protein
A;Residues: 1-89 <GR>
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homo
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinas.
F;6-83/Domain: kringle homology <KRG>
F;6-83/27-66;55-78/Disulfide bonds: #status predicted
F;39/Binding site: carboxydrate (Asn) (covalent) #status experimental

Query Match 61.0%; Score 329; DB 2; Length 89;
Best Local Similarity 61.0%; Pred. No. 2e-25;
Matches 50; Conservative 16; Mismatches 16; Indels 0; Gaps 0;

QY	6	QDCYHGDRSYRGISSTVTGTCQSSMIPHWORTPENYPNAGLTENCRNPDGKQ 65
DB	4	ECYQNGVSYRGTAFTITGKCAWNSMSPHRHNKTESHFNPADLRQNYCRNPDADR 63
QY	66	PWCYTDPQVWEYCNLTQCS 87
DB	64	PWCITDPQVWEYCNLTQCS 85

RESULT 2

S20651
Ig heavy chain V region - mouse
C;Species: Mus musculus (house mouse)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999
C;Accession: S20651; S20647
R;Losman, M.; Fasy, T.M.; Novick, K.E.; Monestier, M.
submitted to the EMBL Data Library, February 1992
A;Description: Relationships among antinuclear antibodies from autoimmune MRL mice r
A;Reference number: S20639
A;Accession: S20651
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-86 <LOS>

A;Cross-references: EMBL:X65007; NID:G52647; PIDN:CAAA6140.1; PID:G52648; EMBL:X6501
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match 12.2%; Score 66; DB 2; Length 86;

Db 26 GSGTSY-----SLTISRMEADAATYQQWSSYPHPVRCWDQAGAE 67
| | | | |
| : : : | : : : | : : : | : : : |

RESULT 10
D96636
unknown protein, 87272-87105 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
C:Accession: D96636
R:rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lib, X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Marzi, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96636
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-55 <STO>
A:CROSS-references: GB:AE005173; NID:g6751695; PIDN:AAF27678.1; GSPDB:GN00141
C:Genetics:
A:Gene: wTP1.20
A:Map position: 1
C:Superfamily: gamma-thionin

Query Match 9.6%; Score 51.5; DB 2; Length 55;
Best Local Similarity 26.6%; Pred No. 1.8e+02;
Matches 17; Conservative 7; Mismatches 15; Indels 25; Gaps 3;

QY 23 TVTGTCQSSSMIPHWHORTENYP-----NAGLTENYCNPDSGKQPWCYT 70
|| ||||| : : : | | | : : : | : : : | : : : |
Db 5 TVEARTCTSSNLF-----NGPCLSSNCANCVHGEFGDGCR---GFRRCLC 51
| | | | |

QY 71 TDPC 74
| | |
Db 52 TRPC 55
| | |

RESULT 11
G69136
ribosomal protein L34 - Methanobacterium thermoautotrophicum (strain Delta H)
N:Alternate names: eukaryotic ribosomal protein L34; prokaryotic ribosomal protein I
C:Species: Methanobacterium thermoautotrophicum
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 13-Aug-1999
C:Accession: G69136
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, K. S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: f
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: G69136
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-88 <XTH>
A:CROSS-references: GB:AE000796; GB:AE000666; NID:g2621057; PIDN:AAB84538.1; PTD:g2621057
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH29
C:Superfamily: rat ribosomal protein L34

Query Match 9.5%; Score 51; DB 2; Length 88;
Best Local Similarity 30.0%; Pred No. 3.1e+02;
Matches 15; Conservative 6; Mismatches 17; Indels 12; Gaps 2;

QY 14 RSYRGISSTVTGTCQSSSMIPHWHTPENYPNAGLTENYCNPDSG 63
| | | | |
| : : : | : : : | : : : | : : : |

Db 9 RSXKXPKVPGGNT-----VTHYRKXKPSKHVCAG-----CKKPLHG 46

RESULT 12

D85557

hypothetical protein 20726 [imported] - Escherichia coli (strain O157:H7, substrain EDL957)

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002

C:Accession: D85557

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, M.W.; Lin, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, N.; et al. Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: D85557

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-72 <STO>

A:Cross-references: GB:AB005174; NID:g12513475; PIDN:AA654920.1; GSPDB:GN00145; UWGP:20726

A:Experimental source: strain O157:H7, substrain EDL953

C:Genetics:

C:Gene: 20726

C:Superfamily: Escherichia coli hypothetical 8K protein (fes 3' region)

Query Match 9.3%; Score 50; DB 2; Length 72;

Best Local Similarity 40.0%; Pred. No. 3.2e+02;

Matches 8; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 28 TCQSWSSMIPHWQRPENY 47

Db 48 SCQQLVY--HWRTLTLPANF 65

RESULT 13

THE2

high potential iron-sulfur protein II - Ectothiorhodospira halophila

N:Alternate names: high-redox-potential ferredoxin 2; HiFip

C:Species: Ectothiorhodospira halophila

C>Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 18-Sep-1998

C:Accession: A00271

R:Tedro, S.M.; Meyer, T.E.; Kamen, M.D.

Arch. Biochem. Biophys. 241, 656-664, 1985

A:Title: Amino acid sequence of high-redox-potential ferredoxin (HiFip) isozymes from the

A:Reference number: A90077; MUID:85305760; PMID:4037807

A:Accession: A00271

A:Molecule type: protein

A:Residues: 1-76 <TED>

C:Comment: The high potential iron-sulfur protein (HiFip) are a class of high-redox-potential ferredoxins found in some other photosynthetic bacteria and in at least one genus (Paracoccus) of halophiles. Two HiFip forms are found.

C:Superfamily: high potential iron-sulfur protein

C:Keywords: 4Fe-4S; electron transfer; iron-sulfur protein; metalloprotein

F:38.41,54.70/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 9.3%; Score 50; DB 1; Length 76;

Best Local Similarity 25.5%; Pred. No. 3.4e+02;

Matches 14; Conservative 6; Mismatches 11; Indels 24; Gaps 3;

QY 29 CQSWSSMIPHWQRPENY 82

Db 41 CQFWVDYVNGW-----GYCQHPD-----FTDVLVRGEGWCSV 72

RESULT 14

JC5892

cobrotoxin b - Chinese cobra

C:Species: Naja naja atra (Chinese cobra)

C>Date: 18-Mar-1998 #sequence_revision 18-Mar-1998 #text_change 21-Jul-2000

C:Accession: JC5892

R:Chang, L.S.; Chou, Y.C.; Lin, S.R.; Wu, B.N.; Lin, J.; Hong, E.; Sun, Y.J.; Hsiao, C.H.; et al. J. Biochem. 122, 1252-1259, 1997

A:Title: A novel neurotoxin, cobrotoxin b, from Naja naja atra (Taiwan cobra) venom: Purification and characterization

A:Reference number: JC5892; MUID:98158338; PMID:9498573

A:Accession: JC5892

A:Molecule type: DNA

A:Residues: 1-82 <CHA>

A:Cross-references: GB:Y13399; NID:g5419941; PIDN:CAA73829.2; PID:g5419942

A:Experimental source: venom

C:Genetics:

A:Introns: 19/1; 54/1

C:Superfamily: snake toxin

Query Match 9.2%; Score 49.5; DB 2; Length 82;

Best Local Similarity 30.0%; Pred. No. 4.1e+02;

Matches 18; Conservative 5; Mismatches 24; Indels 13; Gaps 3;

QY 20 SSTVTGRCQSWSSMIPHW---HQTPEYFNAGLTENYCRNP--DSGKQPCWYTTDPC 74

Db 29 SSQPTTKTCSGETNCKYKKNWSDHRT-----IIRGGCGPKYKPGVNLNCCTTDC 80

RESULT 15

PS0142

replication-associated protein - sugarcane streak virus (fragment)

C:Species: sugarcane streak virus

C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 03-Nov-2000

C:Accession: PS0142

R:Bybicki, E.P.; Hughes, F.L.

submitted to JFIPD, April 1990

A:Description: Detection and typing of maize streak virus and other distantly-related

A:Reference number: PS0142

A:Accession: PS0142

A:Molecule type: DNA

A:Residues: 1-84 <RYB>

C:Genetics:

A:Gene: L1''

C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 9.2%; Score 49.5; DB 2; Length 84;

Best Local Similarity 27.6%; Pred. No. 4.2e+02;

Matches 16; Conservative 4; Mismatches 21; Indels 17; Gaps 3;

QY 8 CYHGDGRSYRGISSTVTGRTCSQSWSSMIPHWQRPENY 65

Db 23 CYDEDA-VYVNIID--IPFKPCQWKLI-----GCQENYVYVNPYKGGK 63

Search completed: September 29, 2003, 08:22:46

Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:04:12 ; Search time 23 Seconds
(without alignments)
186.062 Million cell updates/sec

Title: US-10-088-548-4
Perfect score: 539
Sequence: 1 KSPVVCYHGDSRGIS.....DPCVRWEYCNLTQCSETESG 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 11738

Minimum DB seq length: 0
Maximum DB seq length: 91

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55.5	10.3	86	1 TXM2_DENAN	P18328 dendroaspis
2	55.5	10.3	88	1 CR14_HORVU	P26154 hordeum vul
3	54	10.0	57	1 TXWB_HEMHA	P01402 hemachatus
4	54	10.0	79	1 NXS6_PSETE	O9w177 pseudonaja
5	51	9.5	88	1 RL34_METH	O26137 methanobact
6	50	9.3	76	1 HP12_ECTHA	P04169 ectothiorho
7	49.5	9.2	82	1 CBTB_NAJAT	P80958 naja atra (
8	49	9.1	72	1 YED2_ECOLI	P18393 escherichia
9	49	9.1	81	1 INS_ANAPL	P01333 anas platyr
10	48.5	9.0	72	1 NXLI_DENVI	P01394 dendroaspis
11	48.5	9.0	77	1 THG1_ARATH	O39182 arabidopsis
12	48	8.9	87	1 NXL3_BUNMU	P15817 bungarus mu
13	47.5	8.8	62	1 NXS3_NAJMO	P01432 naja mossam
14	47.5	8.8	77	1 THG2_ARATH	O41914 arabidopsis
15	47	8.7	55	1 ATP8_LATCH	O03168 latimeria c
16	47	8.7	62	1 NXS8_LATLA	P10459 laticauda l
17	46.5	8.6	60	1 TX48_DENJA	P25683 dendroaspis
18	46.5	8.6	77	1 THG3_ARATH	O92ul17 arabidopsis
19	46.5	8.6	80	1 TX3A_PRONI	P81793 phonetria
20	46	8.5	49	1 CBLE_VISAL	P81859 viscum albu
21	46	8.5	61	1 NXS2_HEMHA	P01433 hemachatus
22	46	8.5	63	1 CXH_FEMHA	P24778 hemachatus
23	45.5	8.4	73	1 NXL2_DENVI	P01395 dendroaspis
24	45	8.3	61	1 NXS1_NAJHA	P01429 naja haje a
25	45	8.3	64	1 NXS1_BUNFA	P10808 bungarus fa
26	45	8.3	70	1 R37A_SULSO	O979q3 sulfolobus
27	45	8.3	79	1 NXS1_PSETE	O9w177 pseudonaja
28	44.5	8.3	61	1 CBT4_NAJAT	O97hv0 naja atra (
29	44	8.2	51	1 INS_ANGAN	P07454 anser anser
30	44	8.2	61	1 NXS2_NAJHH	P25675 naja haje h
31	44	8.2	68	1 WT3_BOVIN	P37359 bos taurus
32	44	8.2	76	1 NIKM_BOVIN	O02376 bos taurus
33	44	8.2	78	1 RL28_CORGL	O8ns15 corynebacte

34 44 8.2 79 1 NXS3_PSETE
35 44 8.2 79 1 FLIL_PLETR
36 44 8.2 82 1 CXO6_CONTE
37 43.5 8.1 72 1 NOD1_RHILT
38 43.5 8.1 74 1 EDDF_HUMAN
39 43.5 8.1 87 1 PAGD_SALTY
40 43 8.0 61 1 NXS3_NAJHA
41 43 8.0 61 1 NXS4_NAJHA
42 43 8.0 62 1 SCX2_TITBA
43 43 8.0 62 1 SCX2_TITSE
44 43 8.0 71 1 NXL2_NAJME
45 43 8.0 76 1 TXP4_APTSC

ALIGNMENTS

RESULT 1
TXM2_DENAN STANDARD; PRT; 86 AA.
AC P18328;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Muscarinic toxin 2 precursor.
OS Dendroaspis angusticeps (Eastern green mamba).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Elapidae; Elapinae; Dendroaspis.
OX NCBI_TaxID=8618;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RX MEDLINE=91320365; PubMed=1862524;
RA Ducancel F., Rowan E.G., Cassar E., Harvey A.L., Menez A.,
RA Boullain J.-C.;
RT "Amino acid sequence of a muscarinic toxin deduced from the cDNA
nucleotide sequence."
RL Toxicol 29:516-520(1991).
RN [2]
RP SEQUENCE OF 22-86.
RC TISSUE=Venom;
RX MEDLINE=91320366; PubMed=1862525;
RA Karlsson E., Risinger C., Jolkonen M., Wernstedt C., Adem A.;
RT "Amino acid sequence of a snake venom toxin that binds to the
muscarinic acetylcholine receptor."
RL Toxicol 29:521-526(1991).
CC -!- FUNCTION: BINDS TO THE MUSCARINIC ACETYLCHOLINE RECEPTOR.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: Belongs to the snake toxin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)

CC EMBL; X52292; CAA36541.1; -
CC PIR; A37910; A37910.
DR HSPP; P01382; INTN.
DR InterPro; IPR003572; Cytotoxin.
DR InterPro; IPR003571; Snake_toxin.
DR Pfam; PF00087; toxin; 1.
DR PRINTS; PR00282; CYTOTOXIN.
DR PRODOM; PD000206; Snake_toxin; 1.
DR PROSITE; PS00272; SNAKE_TOXIN; 1.
KW Toxin; Neurotoxin; Signal.
FT SIGNAL 1 21
FT CHAIN 22 86 MUSCARINIC TOXIN 2.
FT DISULFID 24 45 BY SIMILARITY.

```

FT DISULFID 38 63 BY SIMILARITY.
FT DISULFID 67 78 BY SIMILARITY.
FT DISULFID 79 84 BY SIMILARITY.
SQ SEQUENCE 86 AA; 9375 MW; 6F062C970074D653 CRC64;

Query Match 10.3%; Score 55.5; DB 1; Length 86;
Best Local Similarity 28.8%; Pred. No. 27;
Matches 15; Conservative 5; Mismatches 31; Indels 1; Gaps 1;

QY 24 VTGRTCSWSM-IPWHQRTPEPNYPNAGLTENYCRNPDSGKQPCWYTTDPC 74
|| | : : || || || | : | : | || |
Db 33 VTTEDCPAGNVCFKRWVTPKNDIIGKCAATCFKVDNDPIRCCGTDKC 84

RESULT 2
CR14_HORVU
ID CR14_HORVU STANDARD; PRT; 88 AA.
AC P26154;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cold-regulated protein BLT14.
GN BLT14.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Igri; TISSUE=Shoot meristem;
RA Dunn M.A., Hughes M.A., Pearce R.S., Jack P.L.;
RT "Molecular characterization of a barley gene induced by cold
treatment.";
RL J. Exp. Bot. 41:1405-1413(1990).
CC -|- INDUCTION: By cold stress.
CC -|- MISCELLANEOUS: CONTAINS SEVERAL ARGinine RESIDUES IN CLOSE
PROXIMITY WHICH MAY BE INVOLVED IN PROTEIN RNA INTERACTIONS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J57554; CAA40779.1; --
DR PIR; S16161; S16161.
KW RNA-binding.
FT DOMAIN 32 41 ARG-RICH.
SQ SEQUENCE 88 AA; 9796 MW; C125D831E9BB40D0 CRC64;

Query Match 10.3%; Score 55.5; DB 1; Length 88;
Best Local Similarity 29.4%; Pred. No. 28;
Matches 20; Conservative 4; Mismatches 25; Indels 19; Gaps 4;

QY 13 GRSVNGISSTVTGRTCSWSMIPWHQRTPEPNYPNAGLTENYCRNP-----DSGKQPCW 68
|| || || || || | : | : | || |
Db 22 GREBGCSDTRC---RCQRW-----RRRLQFGLAAGGNRYRNKHYPAGGDPW- 69

QY 69 YTTDPCVR 76
||| |
Db 70 ---DPCYR 74

RESULT 3
TXWB_HEMHA
ID TXWB_HEMHA STANDARD; PRT; 57 AA.
AC P01402;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

```

DE Weak toxin CM-1b.
OS Hemachatus haemachatus (Ringhals) (Sapedon haemachatus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Hemachatus.
OX NCBI_TaxID=8626;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=80236878; PubMed=7394807;
RA Joubert F.J., Taljaard N.;
RT "The complete primary structure of toxin CM-1b from Hemachatus
haemachatus (Ringhals) snake venom.";
RL Toxicon 18:191-198(1980).
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -|- MISCELLANEOUS: LD(50) IS 11.7 MG/KG BY INTRAVENOUS INJECTION.
CC -|- SIMILARITY: Belongs to the snake toxin family.
DR PIR; A01673; T2NJB.
DR HSP; P01382; LNTN.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam; PF00087; toxin; 1.
DR ProDom; PD006206; Snake_toxin; 1.
DR PROSITE; PS00272; SNAKE_TOXIN; FALSE_NEG.
KW Toxin.
FT DISULFID 3 19 BY SIMILARITY.
FT DISULFID 12 37 BY SIMILARITY.
FT DISULFID 40 49 BY SIMILARITY.
FT DISULFID 50 55 BY SIMILARITY.
SQ SEQUENCE 57 AA; 6630 MW; 63DIC7818A8B2E20 CRC64;

Query Match 10.0%; Score 54; DB 1; Length 57;
Best Local Similarity 32.7%; Pred. No. 25;
Matches 17; Conservative 7; Mismatches 16; Indels 12; Gaps 4;

QY 28 TCQSWSMTPHHQRTPEPNYPN-----AGLTENYCRNPDSGKQPCWYTTDPC 74
|| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 11 TCQPEQKFC---YSDTMTFFNEHVLSCGT--FCTDSEGER--CCTDRC 55

RESULT 4
NXS6_PSETE
ID NXS6_PSETE STANDARD; PRT; 79 AA.
AC Q9WJ77;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Short neurotoxin 6 precursor (Alpha neurotoxin 6).
GN SNTX6.
OS Pseudonaja textilis (Eastern brown snake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Acanthophiinae; Pseudonaja.
OX NCBI_TaxID=8673;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RX MEDLINE=99449602; PubMed=10518793;
RA Gong N.L., Armugam A., Jeyaseelan K.;
RT "Postsynaptic short-chain neurotoxins from Pseudonaja textilis: cDNA
cloning, expression and protein characterization.";
RL Eur. J. Biochem. 265:982-989(1999).
RN [2]
RP SEQUENCE FROM N.A. AND IDENTIFICATION OF INTRONS.
RC TISSUE=Venom gland;
RX MEDLINE=20279909; PubMed=10818230;
RA Gong N.L., Armugam A., Jeyaseelan K.;
RT "Molecular cloning, characterization and evolution of the genes
encoding a new group of short-chain alpha-neurotoxins in an Australian
elapid, Pseudonaja textilis.";
RL FEBS Lett. 473:303-310(2000).
CC -|- FUNCTION: Lethal neurotoxin, binds and inhibits nicotinic
acetylcholine receptors (nAChR) (By similarity).

```


CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC -----

DR EMEL; AE000796; AAB84538.1; -

DR PIR; G69136; G69136.

DR HAMAP; MF_00349; -; 1.

DR InterPro; IPR001284; Ribosomal_L34E.

DR Pfam; PF011199; Ribosomal_L34e; 1.

DR PRINTS; PR01250; RIBOSOMAL_L34.

DR ProDom; PD005148; Ribosomal_L34E_C; 1.

DR PROSITE; PS01145; RIBOSOMAL_L34E; 1.

DR KQ Ribosomal protein; Complete proteome.

DR KQ SEQUENCE 88 AA; 10301 MW; 3F6C1AC5D2E41ECE CRC64;

DR -----

Query Match 9.5%; Score 51; DB 1; Length 88;

Best Local Similarity 30.0%; Pred. No. 82;

Matches 15; Conservative 6; Mismatches 17; Indels 12; Gaps 2;

DR -----

QY 14 RSYRGISSTVTGRTCSQSSMIPHWHQRTPEYNPAGLTENCRNPDSS 63

DB 9 RSKYRFXKTPGGRT-----VHYRRKKPSKHVCAG-----CGKPLHG 46

DR -----

RESULT 6

HP12_ECTHA STANDARD; PRT; 76 AA.

AC P04169;

ID 20-MAR-1987 (Rel. 04, Created)

DT 20-MAR-1987 (Rel. 04, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE High potential iron-sulfur protein isozyme II (HiPIP 2) (High-redox-potential ferredoxin 2).

DE HIP2.

GN Ectothiorhodospira halophila.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales;

OC Ectothiorhodospiraceae; Halorhodospira.

OX NCBI_TaxID=1053;

XP [1]

RP SEQUENCE.

RX MEDLINE=85305760; PubMed=4037807;

RA Tedro S.M., Meyer T.E., Kamen M.D.;

RT "Amino acid sequence of high-redox-potential ferredoxin (HiPIP) isozymes from the extremely halophilic purple phototrophic bacterium, Ectothiorhodospira halophila.";

RT Arch. Biochem. Biophys. 241:656-664 (1985).

RL -! FUNCTION: SPECIFIC CLASS OF HIGH-REDOX-POTENTIAL 4FE-4S FERREDOXINS. FUNCTIONS IN ANAEROBIC ELECTRON TRANSPORT IN MOST PURPLE AND IN SOME OTHER PHOTOSYNTHETIC BACTERIA AND IN AT LEAST ONE GENUS (PARACOCCLUS) OF HALOPHILIC, DENITRIFYING BACTERIA. THE MIDPOINT REDOX POTENTIAL FOR THIS PROTEIN IS +50 MV.

CC -! SUBUNIT: Homodimer (probable).

CC -! MISCELLANEOUS: IN E. HALOPHILA, TWO HIP1P ISOZYMES ARE FOUND; THEY HAVE THE LOWEST REDOX POTENTIALS YET DETERMINED.

CC -! SIMILARITY: BELONGS TO THE HIGH POTENTIAL IRON-SULFUR PROTEIN (HIP1P) FAMILY.

DR PIR; A0271; IHER2.

DR HSP; P04168; ZHIP.

DR InterPro; IPR000170; Hipot_ironsul.

DR Pfam; PF01355; HIP1P; 1.

DR PRINTS; PR00374; HIP1PFRDOXIN.

DR PROSITE; PS00596; HIP1P; 1.

DR Electron transport; iron-sulfur; 4Fe-4S.

FW METAL 38 38

FT METAL 41 41

FT METAL 54 54

FT METAL 70 70

FT METAL 76 AA; 8560 MW; 12D1524E80BAF7AC CRC64;

DR SEQUENCE

QY Match 9.3%; Score 50; DB 1; Length 76;

Best Local Similarity 25.5%; Pred. No. 89;

Matches 14; Conservative 6; Mismatches 11; Indels 24; Gaps 3;
QY 29 CQSWSSMIPHWQRPENYPNAGLTENYCRNPDGKQPCWYTTDPCYRWE-YCNL 82
DB 41 CQFWDVYNGW-----GYCQHPD-----FTDVLVRGEGWCV 72

RESULT 7
CBTB_NAJAT STANDARD; PRT; 82 AA.
AC P80958; 042285;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cobrotoxin b precursor (CBT b) (Short neurotoxin).
OS Naja atra (Chinese cobra).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Naja.
OX NCBI_TaxID=8656;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 22-82, AND FUNCTION.
RC TISSUE=Liver, and Venom;
RX MEDLINE=96156336; PubMed=9498573;
RA Chang L.-S., Chou Y.-C., Lin S.-R., Wu B.-N., Lin J., Hong E.,
Sun Y.-J., Hsiao C.-D.;
RA "A novel neurotoxin, cobrotoxin b, from Naja naja atra (Taiwan cobra)
RT venom: purification, characterization, and gene organization.";
RL J. Biochem. 122:1252-1259(1997).
RN [2]
RP SEQUENCE FROM N.A., AND REVISION TO 67.
RC TISSUE=Venom gland;
RA Chu R.C., Yang C.C.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Produces peripheral paralysis by blocking neuromuscular
transmission at the postsynaptic site. Binds to the nicotinic
acetylcholine receptor (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: Belongs to the snake toxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; Y13399; CAA73829.2; -;
CC EMBL; AF031472; AAB86636.1; -;
CC PIR; JC5892; JC5892.
CC HSSP; P01427; 1NOR.
CC Pfam; PF00087; Snake_toxin.
CC ProDom; PD000205; Snake_toxin; 1.
CC PROSITE; PS00272; SNAKE_TOXIN; 1.
KW Toxin; Neurotoxin; Postsynaptic neurotoxin; Signal.
FT SIGNAL 1 21
FT CHAIN 22 82 COBROTOXIN B.
FT DISULFID 24 44 BY SIMILARITY.
FT DISULFID 38 61 BY SIMILARITY.
FT DISULFID 63 74 BY SIMILARITY.
FT DISULFID 75 80 BY SIMILARITY.
SQ SEQUENCE 82 AA; 9139 MW; 1FFA21189C08B6E8 CRC64;

Query Match 9.2%; Score 49.5; DB 1; Length 82;
Best Local Similarity 30.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 5; Mismatches 24; Indels 13; Gaps 3;
QY 20 SSTVTGRTQCSWSSMIPHW-----HQTPEYFNAGLTENYCRNP--DSGKQPCWYTTDPC 74
DB 29 SSTPTTKTCGTNCYKKNWSDHGT-----IIERGCGCPKPKGVNLNCCITDRC 80

RESULT 8
YBDZ_ECOLI STANDARD; PRT; 72 AA.
AC P18393;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ybdz.
GN YBDZ OR B0585.1.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89066678; PubMed=2974033;
RA Pettis G.S., Brickman T.J., McIntosh M.A.;
RT "Transcriptional mapping and nucleotide sequence of the Escherichia
coli fepA-fes enterobactin region. Identification of a unique
iron-regulated bidirectional promoter.";
RL J. Biol. Chem. 263:18857-18863(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91175738; PubMed=1826089;
RA Rusnak F., Sakaitani M., Drucehammer D., Reichert J., Walsh C.T.;
RT "Biosynthesis of the Escherichia coli siderophore enterobactin:
sequence of the entf gene, expression and purification of EntP, and
analysis of covalent phosphopantetheine.";
RL Biochemistry 30:2916-2927(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -!- FUNCTION: COULD BE INVOLVED IN ENTEROBACTIN SYNTHESIS.
CC -!- SIMILARITY: TO M.TUBERCULOSIS METH.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; J04216; AAA23758.1; -;
CC EMBL; M60177; -; NOT_ANNOTATED_CDS.
CC PIR; AE000163; -; NOT_ANNOTATED_CDS.
CC PIR; B31958; Q3ECFS.
CC EcoGene; EG14353; ybdz.
DR InterPro; IPR005153; Mbth.
DR Pfam; PF03621; Mbth; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 72 AA; 8271 MW; 164E0B87E00F1BB8 CRC64;

Query Match 9.1%; Score 49; DB 1; Length 72;
Best Local Similarity 40.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 3; Mismatches 7; Indels 2; Gaps 1;
QY 28 TCQSWSSMIPHWQRPENY 47
DB 48 SCQQW--LEAHWRTLTPTNF 65

RESULT 9
INS_ANAPL

```

ID  INS_ANAPL  STANDARD;  PRT;  81 AA.
AC  P01333;
DT  21-JUL-1986 (Rel. 01, Created)
DT  21-JUL-1986 (Rel. 01, Last sequence update)
DT  15-JUL-1999 (Rel. 38, Last annotation update)
DE  Insulin precursor.
GN  INS.
OS  Anas platyrhynchos (Domestic duck).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
OX  NCBI_TaxID=8839;
RN  [1]
RP  SEQUENCE OF 1-30 AND 61-81.
RX  MEDLINE=74055140; PubMed=4763354;
RA  Markussen J., Sundby F.;
RT  "Duck insulin: isolation, crystallization and amino acid sequence.";
RL  Int. J. Pept. Protein Res. 5:137-48(1973).
RN  [2]
RP  SEQUENCE OF 33-58.
RX  MEDLINE=73210102; PubMed=4715652;
RA  Markussen J., Sundby F.;
RT  "Isolation and amino-acid sequence of the C-peptide of duck
    proinsulin.";
RL  Eur. J. Biochem. 34:401-408(1973).
CC  -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
    INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
    FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
    CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC  -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
    DISULFIDE BONDS.
CC  -1- SUBCELLULAR LOCATION: Secreted.
CC  -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC  -1- CAUTION: X'S AT POSITIONS 31-32 AND 59-60 REPRESENT PAIRED BASIC
    RESIDUES ASSUMED BY HOMOLOGY TO BE PRESENT IN THE PRECURSOR
    MOLECULE.
DR  PIR; A01600; IPDK.
DR  HSP; P01308; IHIS.
DR  InterPro; IPR004825; Ins/IGF/relax.
DR  Pfam; PF00049; Insulin; 1.
DR  SMART; SM00078; ILGF; 1.
DR  PROSITE; PS00262; INSULIN; 1.
KW  Insulin family; Hormone; Glucose metabolism.
FT  CHAIN 1 30 INSULIN B CHAIN.
FT  PROPEP 33 58 C PEPTIDE.
FT  CHAIN 61 81 INSULIN A CHAIN.
FT  DISULFID 7 67 INTERCHAIN.
FT  DISULFID 19 80 INTERCHAIN.
FT  DISULFID 66 71
SQ  SEQUENCE 81 AA; 9100 MW; 6E8A271F099DA91 CRC64;

Query Match 9.1%; Score 49; DB 1; Length 81;
Best Local Similarity 30.4%; Pred. No. 1.2e+02;
Matches 14; Conservative 3; Mismatches 15; Indels 14; Gaps 3;

QY  36 IPHWQRTPEPNYAGLTENYCNPDGKQPCWYTTDPCVRWEYCN 81
    :| | | | | | | | | | | | | | | | | | | |
DB  50 LPFQF----EYQXGIVECCNPNCS-----LYQLE-----NYCN 81
    :| | | | | | | | | | | | | | | | | | | |

RESULT 10
NALLI_DENVI
ID  NALLI_DENVI  STANDARD;  PRT;  72 AA.
AC  P01394;
DT  21-JUL-1986 (Rel. 01, Created)
DT  21-JUL-1986 (Rel. 01, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Long neurotoxin 1 (Neurotoxin 4.7.3/4.9.3).
OS  Dendroaspis viridis (Neurotoxin green mamba).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC  Elapidae; Elapinae; Dendroaspis.
OX  NCBI_TaxID=8621;
RN  [1]

```

```

RP  SEQUENCE.
RX  TISSUE=Venom;
RX  MEDLINE=74309059; PubMed=4855008;
RA  Banks B.E.C., Miledi R., Shipollini R.A.;
RT  "The primary sequences and neuromuscular effects of three neurotoxic
    polypeptides from the venom of Dendroaspis viridis.";
RL  Eur. J. Biochem. 45:457-468(1974).
CC  -1- SUBCELLULAR LOCATION: Secreted.
CC  -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC  -1- PTM: NEUROTOXIN 4.7.3 DIFFERS FROM 4.9.3 ONLY IN THAT
    TRP-26 HAS UNDERGONE PARTIAL PHOTOXIDATION.
CC  -1- MISCELLANEOUS: LD(50) IS 0.9 MG/KG BY INTRAPERITONEAL INJECTION.
CC  -1- SIMILARITY: Belongs to the snake toxin family.
DR  PIR; A01665; N2EP1V.
DR  HSP; P01386; 1FXA.
DR  InterPro; IPR003571; Snake_toxin.
DR  Pfam; PF00087; toxin; 1.
DR  ProDom; PD000206; Snake toxin; 1.
DR  PROSITE; PS00272; SNAKE_TOXIN; 1.
KW  Toxin; Neurotoxin; Multigene family.
FT  DISULFID 3 21 BY SIMILARITY.
FT  DISULFID 14 42 BY SIMILARITY.
FT  DISULFID 27 31 BY SIMILARITY.
FT  DISULFID 46 57 BY SIMILARITY.
FT  DISULFID 58 63 BY SIMILARITY.
SQ  SEQUENCE 72 AA; 7970 MW; F299DA0782163B4F CRC64;

Query Match 9.0%; Score 48.5; DB 1; Length 72;
Best Local Similarity 26.3%; Pred. No. 1.2e+02;
Matches 20; Conservative 6; Mismatches 31; Indels 19; Gaps 4;

QY  1 KSPVY--QDCVHGGRSYRISSTTGTGTCQSSMIPHWHORTPENYNAGLTENYCR 58
    :| | | | | | | | | | | | | | | | | | | |
DB  5 KIPSVKPTCTCHGENICY-----TEWCDANCS-----QRGKREELGCAAT---CP 47
    :| | | | | | | | | | | | | | | | | | | |

QY  59 NPDGKQPCWYTTDPC 74
    :| | | | | | | | | | | | | | | | | | | |
DB  48 KYKAGVGIKCCSTDC 63

RESULT 11
THGLI_ARATH
ID  THGLI_ARATH  STANDARD;  PRT;  77 AA.
AC  Q39182; Q42011;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Gamma-thionin homolog At2G02100 precursor.
GN  AT2G02100 OR F504.13.
OS  Arabidopsis thaliana (Mouse-ear cress).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC  eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX  NCBI_TaxID=3702;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=cv. C24; TISSUE=Flower buds;
RA  Yu D.Y., Quigley F., Mache R.;
RT  "Isolation and expression of a cDNA encoding protease inhibitor
    precursor.";
RL  Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=cv. Columbia;
RA  Lin X., Kaul S., Rounsley S.D., Shear T.P., Benito M.-I., Town C.D.,
RA  Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA  Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA  Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA  Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA  Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA  Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA  Venter J.C.;

```


DR	Pfam: PF00895; ANP-synt_8; 1.
KW	Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT	SEQUENCE 4 24 POTENTIAL.
SQ	SEQUENCE 55 AA; 6569 MW; 7FE36319E8AF825B CRC64;

Search completed: September 29, 2003, 08:20:19
Job time : 25 secs

DB 64 CHTINSQVRWEYCKTPSCS 82

Search completed: September 29, 2003, 08:19:47
Job time : 83 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:11:29 ; Search time 93 Seconds
(without alignments)
252.503 Million cell updates/sec

Title: US-10-088-548-4
Perfect score: 539
Sequence: 1 KSPVQDCYHGDGRSGRIS.....DPCVRWEYNLTQCSFESG 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 118357

Minimum DB seq length: 0
Maximum DB seq length: 91

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTRMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_prodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	236	43.8	53	Q8UMI2	Q8UMI2 homo sapien
2	207.5	38.5	90	Q8NG20	Q8NG20 homo sapien
3	188	34.9	60	Q9UKJ7	Q9UKJ7 homo sapien
4	120.5	22.4	75	Q9BGN9	Q9BGN9 bos taurus
5	109	20.2	25	Q9UD88	Q9UD88 homo sapien
6	62	11.5	69	Q8MVA6	Q8MVA6 ixodes scap
7	60.5	11.2	87	Q8V5I3	Q8V5I3 human respi
8	59.5	11.0	82	Q30599	Q30599 macaca mula
9	59.5	11.0	82	Q19292	Q19292 macaca mula
10	59	10.9	78	Q8SQ84	Q8SQ84 gorilla gor
11	58.5	10.9	87	Q8BBA5	Q8BBA5 human respi
12	58.5	10.9	87	Q8BBA4	Q8BBA4 human respi
13	58	10.8	64	Q3DF51	Q3DF51 oncorhynchus
14	57.5	10.7	77	Q39403	Q39403 brassica ca
15	57.5	10.7	83	Q8V5I2	Q8V5I2 human respi
16	57.5	10.7	83	Q8BBA8	Q8BBA8 human respi

17	57.5	10.7	87	12	Q8V3I4	Q8V3I4 human respi
18	57	10.6	85	11	Q9JHV3	Q9JHV3 mus musculu
19	56.5	10.5	80	12	Q8V5I6	Q8V5I6 human respi
20	56.5	10.5	80	12	Q9IUD1	Q9IUD1 human respi
21	56.5	10.5	83	12	Q9IUE3	Q9IUE3 human respi
22	56.5	10.5	83	12	Q9IH96	Q9IH96 human respi
23	56.5	10.5	83	12	Q8V5H5	Q8V5H5 human respi
24	56.5	10.5	83	12	Q8BBA3	Q8BBA3 human respi
25	56.5	10.5	87	12	Q8V5I7	Q8V5I7 human respi
26	55.5	10.3	65	13	Q9PRY3	Q9PRY3 dendroaspis
27	54	10.0	72	11	Q99LN5	Q99LN5 mus musculu
28	53.5	9.9	80	12	Q9IH93	Q9IH93 human respi
29	53.5	9.9	81	7	Q30728	Q30728 macaca neme
30	53.5	9.9	83	12	Q9IH91	Q9IH91 human respi
31	53.5	9.9	83	12	Q9IH88	Q9IH88 human respi
32	53.5	9.9	83	12	Q9IH92	Q9IH92 human respi
33	53.5	9.9	83	12	Q9IH87	Q9IH87 human respi
34	53.5	9.9	83	12	Q9IUD2	Q9IUD2 human respi
35	53.5	9.9	83	12	Q9IMX7	Q9IMX7 human respi
36	53.5	9.9	83	12	Q9IH95	Q9IH95 human respi
37	53.5	9.9	83	12	Q9IUC9	Q9IUC9 human respi
38	53	9.8	70	7	Q46870	Q46870 homo sapien
39	53	9.8	85	6	Q46556	Q46556 equus cabal
40	52.5	9.7	66	4	Q9H409	Q9H409 homo sapien
41	52.5	9.7	81	7	Q8WLN3	Q8WLN3 papio anubi
42	52	9.6	76	4	Q16198	Q16198 homo sapien
43	51.5	9.6	55	10	Q9C947	Q9C947 arabidopsis
44	51.5	9.6	80	7	Q8XR03	Q8XR03 salmo salar
45	51.5	9.6	82	7	Q9MXN5	Q9MXN5 macaca mula

ALIGNMENTS

RESULT 1

ID	Q9UMI2	PRELIMINARY;	PRT;	53 AA.
AC	Q9UMI2;			
DT	01-MAY-2000 (TrEMBLrel. 13, Created)			
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)			
DE	PLG protein (fragment).			
GN	PLG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85023311; PubMed=6148961;			
RA	Mallinowski D.P., Sadler J.E., Davie E.W.;			
RT	"Characterization of a complementary deoxyribonucleic acid coding for			
RL	human and bovine plasminogen."			
CC	-!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.			
DR	EMBL; K02921; AAA60123.1; -			
DR	HSSP; P00747; 2PK4			
DR	InterPro; IPR000001; Kringle.			
DR	Ffam; PF00051; Kringle; 1.			
DR	PRINTS; PR00018; KRINGLE.			
DR	ProDom; PD000395; Kringle; 1.			
DR	SMART; SM00130; KR; 1.			
DR	PROSITE; PS00021; KRINGLE_1; FALSE_NEG.			
DR	PROSITE; PS50070; KRINGLE_2; 1.			
KW	Glycoprotein; Kringle.			
FT	NON_TER 1			
SQ	SEQUENCE 53 AA; 5893 MW; 2EB36E715CD40A47 CRC64;			
Query Match		43.8%;	Score 236;	DB 4; Length 53;
Best Local Similarity		81.8%;	Pred. No. 3.4e-19;	
Matches 40;	Conservative	4;	Mismatches 5;	Indels 0;
Gaps				0;
QY	2 SPVQDCYHGDGRSGRISSTVTGTGTCQSSSMIPHWHTPENYPNA 50			


```

Db 5 TPVQDYGQSGVSGTSTTTTGKCKQSSWTPPHHQKTPENYPNA 53
      ::::::::::::::::::::: ||| ||| :::::::::: || ::::::::::
RESULT 2
Q8NG20 PRELIMINARY; PRT; 90 AA.
AC Q8NG20;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Plasminogen/activator kringle.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dou D.;
RT "Production of kringle fragment.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF282882; AAM52248.1; -.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PSS0070; KRINGLE_2; 1.
KW Glycoprotein; Kringle.
SQ SEQUENCE 90 AA; 9804 MW; A33887F9FDF4C7B1 CRC64;

Query Match 38.5%; Score 207.5; DB 4; Length 90;
Best Local Similarity 47.6%; Pred. No. 9.2e-16;
Matches 40; Conservative 14; Mismatches 25; Indels 5; Gaps 3;

QY 7 DCHGDRYRGISSTVTTGRTQSSMIPHWHTQTPENYPNA---GL-TENYCRNPD 62
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7 DCFYNGSAYRGTSLTESGASCLPWNMLIGKVTYQN-PSAALGLGKHYCRNPDG 65
      ::::::::::: ||| ||| ::::::::::

QY 63 GKQPCWYTDPCVRWEYCNLTQCS 86
      ::::::::::: ||| ||| ::::::::::
Db 66 DAKPCWYTTNPKLYDYDVPQCA 89
      ::::::::::: ||| ||| ::::::::::

RESULT 3
Q9UKJ7 PRELIMINARY; PRT; 60 AA.
AC Q9UKJ7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21181705; PubMed=11285247;
RA Ogorekova M., Kraft H.G., Enholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158657; AAF03676.1; -.
DR HSP; P00747; LKRN.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRODOM; PD000395; Kringle; 1.

```

```

DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PSS0070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoaprotein.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 6799 MW; 5719AA26B3EOFFLD CRC64;

Query Match 34.9%; Score 188; DB 4; Length 60;
Best Local Similarity 77.5%; Pred. No. 9e-14;
Matches 31; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 51 GLTENYCRNPDGKQPCWYTDPCVRWEYCNLTQCSSETS 90
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 GLTRYCRNPDAEIRPWCYTMDPSVWEYCNLTQCLYTES 40
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
Q9BGN9 PRELIMINARY; PRT; 75 AA.
AC Q9BGN9;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hepatocyte growth factor (fragment).
GN HGF.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=uterus;
RA Murakami S., Fujiwara C., Miyamoto Y., Takeuchi S., Takahashi S.,
RA Okuda K.;
RT "Expression and action of hepatocyte growth factor in bovine
RT endometrial stromal and epithelial cells in vitro.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AB056447; BAB33031.1; -.
DR HSP; P14210; IBHT.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 2.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PSS0070; KRINGLE_2; 2.
KW Glycoprotein; Kringle.
FT NON_TER 1
FT NON_TER 75
SQ SEQUENCE 75 AA; 8831 MW; 829EEEFCC49701B1 CRC64;

Query Match 22.4%; Score 120.5; DB 6; Length 75;
Best Local Similarity 56.8%; Pred. No. 4e-06;
Matches 21; Conservative 8; Mismatches 5; Indels 3; Gaps 2;

QY 55 NYCRRNPDSGKQ--PWCYTDPQVRWEYCNLTQCSSETS 89
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 NYCRRNP-RGEGGPGWCFTSNPEVRYEVCDDIPQCSSEVE 36
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
Q9UD88 PRELIMINARY; PRT; 25 AA.
AC Q9UD88;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE APO(A) KRINGLE 4-37, APO(A) KRINGLE type 5 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94060120; PubMed=7848387;
RA Praeflinger D., McLean J., Scanu A.M.;
RT "Amplification of human APO(a) kringle 4-37 from blood lymphocyte
  DNA.";
RL Biochim. Biophys. Acta 1225:107-109(1993).
DR HSP: P00747; 1KEN
DR InterPro: IPR000001; Kringle.
DR Pfam: PF00051; Kringle.1.
KW Glycoprotein; Kringle.
FT NON_TER 1 1
FT TER 25 25
SQ SEQUENCE 25 AA; 2928 MW; BEED4C62FA480A8 CRC64;

Query Match 20.2%; Score 109; DB 4; Length 25;
Best Local Similarity 72.7%; Pred. No. 2.3e-05;
Matches 16; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 68 CYTDPVWEYCNLTQCSSE 89
Db 1 CFTMDPSIRWEYCNLTQCSDE 22

RESULT 6
Q8MVA6 PRELIMINARY; PRT; 69 AA.
AC Q8MVA6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Putative 5.3 kDa secreted protein.
OS Ixodes scapularis (Black-legged tick) (Deer tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Ixodes.
OX NCBI_TaxID=6945;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rhode Island; TISSUE=Salivary gland;
RA Valenzuela J.G., Francischetti I.N., Pham V.M., Garfield M.,
RA Mather T.N., Ribeiro J.M.C.;
RT "Exploring the Sialome of the Tick Vector of Lyme Disease, Ixodes
  scapularis.";
RL Submitted (FBI-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF483734; AAK93656.1;
SQ SEQUENCE 69 AA; 7817 MW; EBB93FDBFA754E24 CRC64;

Query Match 11.5%; Score 62; DB 5; Length 69;
Best Local Similarity 35.0%; Pred. No. 13;
Matches 14; Conservative 5; Mismatches 11; Indels 10; Gaps 3;

QY 60 PDSCKQPW-----CYTDPVWEYCNLTQCSSE 91
Db 24 PDGP-QPWQVQKAGRPFCYSI-PCRKHDECRVGCSCRNNG 61

RESULT 7
Q8V5I3 PRELIMINARY; PRT; 87 AA.
AC Q8V5I3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Attachment glycoprotein G (Fragment).
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-SA98VI53;
RX MEDLINE=21405813; PubMed=11514720;
RA Venter M., Madhi S.A., Tiemessen C.F., Schoub B.D.;
RT "Genetic diversity and molecular epidemiology of respiratory syncytial
  virus over four consecutive seasons in South Africa: identification of
  new subgroup A and B genotypes.";
RL J. Gen. Virol. 82:2117-2124(2001).
DR EMBL: AF348815; AAL60316.1;
DR InterPro: IPR000925; Glycoprot_G.
DR Pfam: PF00802; Glycoprotein_G; 1.
FT NON_TER 1 1
FT TER 87 87
SQ SEQUENCE 87 AA; 9431 MW; CEF8B4DD27344769 CRC64;

Query Match 11.2%; Score 60.5; DB 12; Length 87;
Best Local Similarity 32.1%; Pred. No. 24;
Matches 17; Conservative 4; Mismatches 27; Indels 5; Gaps 1;

QY 1 KSPVQDCYHGDSYRGISSTVTGRTQCSWSSMIPHWHQRTPENYPNAGIT 53
Db 21 KKPTPKTERTGTSQSSTVLDTTTSKHTIQQS-----LHSITPTNPSTQT 68

RESULT 8
Q30599 PRELIMINARY; PRT; 82 AA.
AC Q30599;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Major histocompatibility complex class II (Fragment).
GN MHC-DRB.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97152405; PubMed=8995183;
RA Knapp J.A., Cadavid L.F., Eberle M.E., Knechtle S.J., Bontrop R.E.,
RA Watkins D.I.;
RT "Identification of new manu-DRB alleles using DGGE and direct
  sequencing.";
RL Immunogenetics 45:171-179(1997).
DR EMBL: U57941; AAC50974.1;
DR InterPro: IPR000353; MHC_II_beta.
DR Pfam: PF00969; MHC_II_beta; 1.
DR ProDom: PD000328; MHC_II_beta; 1.
KW Glycoprotein; MHC II; Transmembrane.
FT NON_TER 1 1
FT TER 82 82
SQ SEQUENCE 82 AA; 9945 MW; 1FE11A6520A60108 CRC64;

Query Match 11.0%; Score 59.5; DB 7; Length 82;
Best Local Similarity 30.8%; Pred. No. 29;
Matches 16; Conservative 8; Mismatches 17; Indels 11; Gaps 3;

QY 9 YEGDGRSYRGISSTVTGRTQCSWSSMIPHWHQRTPENYPNAGLTENYCRN 59
Db 33 FDSVDVGEYRVSF---LGRPTAESWNSQKDYLEQ-----RGQVDNYCRH 74

RESULT 9
Q19292 PRELIMINARY; PRT; 82 AA.
AC Q19292;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MHC class II DR beta (fragment).
GN HLA-DRB.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE OF 3-30 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=94179810; PubMed=8133043;
RA Sliendregt B.L., Oetting N., van Besouw N., Jonker M., Bontrop R.E.;
RT "Expansion and contraction of rhesus macaque DRB regions by
RT duplication and deletion.";
RL J. Immunol. 152:2298-2307(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA de Groot N.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF031268; AB87440.1; -.
DR HSSP; P13758; IDLH.
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
KW Glycoprotein; MHC II; Transmembrane.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 9960 MW; 71FA3A6520A60119 CRC64;

Query Match 11.0%; Score 59.5; DB 7; Length 82;
Best Local Similarity 30.8%; Pred. No. 29;
Matches 16; Conservative 8; Mismatches 17; Indels 11; Gaps 3;

QY 9 YHGDSYRGISSTVTGCR-TCQSWSSMIPHWQRTPEPNAGLTENYCRN 59
DB 33 FDSVDGEYRAVSE--LGRPTAESWNSQDYLEQKRAE-----VDNYCRH 74

RESULT 10
Q8SQ84 ID Q8SQ84 PRELIMINARY; PRT; 78 AA.
AC Q8SQ84;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE MHC-G (Fragment).
GN MHC-G.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=97364681; PubMed=9218527;
RA Arnaiz-Villena A., Martinez-Laso J., Alvarez M., Castro M.J.,
RA Varela P., Gomez-Casado E., Suarez B., Recio M.J., Vargas-Alarcon G.,
RA Morales P.;
RT "Primate Mhc-B and -G alleles.";
RL Immunogenetics 46:251-266(1997).
DR EMBL; I41256; AAL77572.1; -.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00129; MHC_I; 1.
DR ProDom; PD000050; MHC_I; 1.
FT NON_TER 1
FT NON_TER 78
SQ SEQUENCE 78 AA; 8891 MW; EC12D3F3A58FC7B3 CRC64;

Query Match 10.9%; Score 59; DB 6; Length 78;
Best Local Similarity 22.7%; Pred. No. 31;
Matches 15; Conservative 10; Mismatches 37; Indels 4; Gaps 1;

QY 12 DGRSYRGISSTVTGRTQCSWSSMIPHWQRTPEPNAGLTENYCRNPSGKOPWCYTT 71
DB 7 DGRLLRGYDGYADGKDYLALNEDLRSWTAADTA----AQISKRCCEAPNSQRRAYLE 62

QY 72 DPCVRW 77
DB 63 GTCVEW 68

RESULT 11
Q8BBA5 ID Q8BBA5 PRELIMINARY; PRT; 87 AA.
AC Q8BBA5;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Glycoprotein (Fragment).
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AGK6-99;
RX MEDLINE=22213634; PubMed=12226836;
RA Venter M., Collinson M., Schoub B.D.;
RT "Molecular epidemiological analysis of community circulating
RT respiratory syncytial virus in rural South Africa: Comparison of
RT viruses and genotypes responsible for different disease
RT manifestations.";
RL J. Med. Virol. 68:452-461(2002).
DR EMBL; AF548806; AAN62463.1; -.
FT NON_TER 1
FT NON_TER 87
SQ SEQUENCE 87 AA; 9431 MW; F72B9DC59C975EE0 CRC64;

Query Match 10.9%; Score 58.5; DB 12; Length 87;
Best Local Similarity 32.7%; Pred. No. 40;
Matches 16; Conservative 3; Mismatches 25; Indels 5; Gaps 1;

QY 1 KSPVQDCYHGDGRSYRGISSTVTGRTQCSWSSMIPHWQRTPEPNYPN 49
DB 21 KKPTKTRGTSTQSTVLDTTTSHKTIQQQS-----LHSTTPENTPN 64

RESULT 12
Q8BBA4 ID Q8BBA4 PRELIMINARY; PRT; 87 AA.
AC Q8BBA4;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Glycoprotein (Fragment).
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AGU16.99;
RX MEDLINE=22213634; PubMed=12226836;
RA Venter M., Collinson M., Schoub B.D.;
RT "Molecular epidemiological analysis of community circulating
RT respiratory syncytial virus in rural South Africa: Comparison of
RT viruses and genotypes responsible for different disease
RT manifestations.";
RL J. Med. Virol. 68:452-461(2002).
DR EMBL; AF548807; AAN62464.1; -.
FT NON_TER 1
FT NON_TER 87
SQ SEQUENCE 87 AA; 9491 MW; 3BFB3AD57345C68 CRC64;

Query Match 10.9%; Score 58.5; DB 12; Length 87;
Best Local Similarity 32.7%; Pred. No. 40;
Matches 16; Conservative 3; Mismatches 25; Indels 5; Gaps 1;

QY 1 KSPVQDCYHGDGRSYRGISSTVTGRTQCSWSSMIPHWQRTPEPNYPN 49

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:22:53 ; Search time 82 Seconds
(without alignments)
172.276 Million cell updates/sec

Title: US-10-088-548-6
Perfect score: 523
Sequence: 1 VRQYHGNGQSYRGTFSTV.....SIRWEYCNLTRCSDTEGTVV 89

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 6931114

```
Minimum DB seq length: 0
Maximum DB seq length: 89
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

- 1: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSL/cgcdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB	ID	Description
1	523	100.0	89	22	AAV72946		Human apolipoprotein
2	430	82.2	84	21	AAV77720		Human plasminogen
3	427	81.6	82	20	AAP93231		Plasminogen kringle 1
4	421	80.5	83	21	AAH08414		Amino acid sequence
5	417	79.7	78	17	AAW07560		Human kringle 4.
6	369	70.6	78	17	AAW07559		Murine kringle 4.
7	349.5	66.8	86	23	AAO32801		Human ORFX protein
8	313	59.8	78	17	AAW07554		Murine kringle 3.
9	299	57.2	78	17	AAW07557		Porcine kringle 3.

10	286	54.7	83	21	AAB01908	Human plasminogen
11	285	54.5	78	17	AAW07555	Human kringle 3.
12	283	54.1	78	17	AAW07556	Rhesus kringle 3.
13	281	53.7	78	17	AAW07558	Bovine kringle 3.
14	277	53.0	78	17	AAW07553	Bovine kringle 2.
15	276	52.8	83	21	AAB08442	Amino acid sequence
16	274	52.4	78	17	AAW07552	Porcine kringle 2.
17	271.5	51.9	89	20	AAW02134	SBQ ID 113 of W099
18	271.5	51.9	89	20	AAW02133	SBQ ID 113 of W099
19	271.5	51.9	89	20	AAW02110	A multifunctional
20	269.5	51.5	82	21	AAW77719	Human plasminogen
21	269.5	51.5	84	10	AAW95230	plasminogen kringle
22	269.5	51.5	84	21	AAB01906	Human plasminogen
23	268.5	51.3	79	17	AAW07546	Rhesus kringle 1.
24	267.5	51.1	79	17	AAW07545	human kringle 1.
25	264	50.5	78	17	AAW07550	Human kringle 2.
26	264	50.5	85	21	AAW01907	Human plasminogen
27	262.5	50.2	75	9	AAW81360	K1 domain of plasmin
28	262.5	50.2	79	17	AAW96221	Plasminogen kringle
29	262	50.1	78	17	AAW07551	Rhesus kringle 2.
30	261	49.9	78	21	AAW77721	Human plasminogen
31	260	49.7	78	17	AAW07549	Murine kringle 2.
32	256.5	49.0	79	17	AAW07548	Bovine kringle 1.
33	256	48.9	81	10	AAW90598	Synthetic N-termin
34	254.5	48.7	78	10	AAW94401	Sequence encoding
35	253.5	48.5	84	21	AAW08440	Amino acid sequence
36	247.5	47.3	79	17	AAW07547	Porcine kringle 1.
37	245.5	46.9	79	17	AAW07544	murine kringle 1.
38	244	46.7	79	18	AAW19256	Human plasminogen
39	240	45.9	85	21	AAW08443	Amino acid sequence
40	236	45.1	82	21	AAW08441	Amino acid sequence
41	224	42.8	86	22	AAW72947	Human apolipoprote
42	223	42.6	52	22	ABG51665	Human liver peptid
43	223	42.6	52	22	ABG31599	Peptide #4250 enco
44	223	42.6	52	22	ABG36820	Peptide #4326 enco
45	223	42.6	52	22	ABB22143	Protein #442 enco

ALIGNMENTS

RESULT 1

AAW72946

ID AAW72946 standard; Protein; 89 AA.

XX AC AAW72946;

XX AC AAW72946;

DT 13-JUN-2001 (first entry)

XX XX

DE Human apolipoprotein(a) kringle domain IV37, LK7 protein.

XX XX

KW Human; angiogenesis inhibitor; LK7; apolipoprotein(a) kringle domain;

KW angiogenesis-mediated disease; cancer; rheumatoid arthritis; therapy;

KW cytostatic; antirheumatic; antiarthritic; antipsoriatic; psoriasis;

KW ocular angiogenic disease; endothelial cell proliferation; tumour;

cell migration.

XX XX

OS Homo sapiens.

XX XX

PN WO200119868-A1.

XX XX

PD 22-MAR-2001.

XX XX

PF 15-SEP-1999; 99WO-KR00554.

XX XX

PR 15-SEP-1999; 99WO-KR00554.

XX XX

PA (MOGA-) MOGAM BIOTECHNOLOGY RES INST.

XX XX

PI Chang J, Kim JS, Park EJ, Yum J, Chung S;

XX XX

DR WPI; 2001-244787/25.

DR N-PSDB; AAD03257.

```

XX Novel angiogenesis inhibitor, LK68 for treating angiogenesis-mediated
PT diseases, e.g. cancer and rheumatoid arthritis, has human
PT apolipoprotein(a) kringle domains IV36, IV37 and V38 amino acid
PT sequence
XX
PS Claim 2; Page 45-46; 50pp; English.
XX
CC The present sequence is human LK7 protein which contains the amino
CC acid sequence of human apolipoprotein(a) kringle domain IV37
CC (KIV37). The human apolipoprotein(a) kringle domains IV36 (LK6
CC protein), IV37 (LK7 protein) and V38 (LK8 protein) together form the
CC angiogenesis inhibitor, LK68 protein. LK68, LK6, LK7 and LK8 are
CC inhibitors are of endothelial cell proliferation, cell migration and
CC normal development of capillaries in the chick embryo chorioallantoic
CC membrane (CAM). LK68 protein, its single kringle or their functional
CC equivalents, are useful for treating angiogenesis-mediated diseases,
CC such as cancer, rheumatoid arthritis, psoriasis or ocular angiogenic
CC disease in animals or humans. LK68 is useful as an anticancer agent and
CC also for inhibiting primary tumour growth.
XX
SQ Sequence 89 AA;
Query Match 100.0%; Score 523; DB 22; Length 89;
Best Local Similarity 100.0%; Pred. No. 5.2e-46;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VRQYHGNGQSYRGTFSTVTGTCQSSMTPHRHQTPEPNDGLTMNCRNPDADT 60
DB 1 VRQYHGNGQSYRGTFSTVTGTCQSSMTPHRHQTPEPNDGLTMNCRNPDADT 60
QY 61 GPWCFTTDPDSIRWEYCNLTRCSDTGTGVV 89
DB 61 GPWCFTTDPDSIRWEYCNLTRCSDTGTGVV 89
RESULT 2
RAY77720
ID AAY77720 standard; protein: 84 AA.
AC AAY77720;
XX
XX 12-MAY-2000 (first entry)
DE Human plasminogen kringle 4 sequence.
KW Kringle 4; plasminogen; anti-angiogenic; angiogenesis; angiotensin;
KW cystostatic; antiarthritic; antirheumatic; antidiabetic; ophthalmological;
KW immunosuppressant; vasotropic; vulnerary; antiarteriosclerotic; human;
KW dermatological; cancer; tumour; birth control; vascularization.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT Region 59..75
FT /note= "specifically claimed fragment (AAY77723)"
XX
XX W0200003726-A1.
XX
XX 27-JAN-2000.
XX
XX 07-JUL-1999; 99WO-US15271.
XX
XX 14-JUL-1998; 98US-0092831.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Ji R, Trail PA;
XX
XX WPI; 2000-171200/15.
XX
XX Novel lysine binding fragments angiotensin used as antiangiogenic
PT agents in the treatment of cancer, diabetic retinopathy, rheumatoid

```

```

PT arthritis, psoriasis, atherosclerotic plaque formation, and other
PT angiogenesis diseases
XX
PS Disclosure; Fig 3; 30pp; English.
XX
CC The invention provides fragments of kringle 1, 2 or 4 of human
CC plasminogen that contain a lysine binding site and have anti-angiogenic
CC activity. The peptides of the invention function as antiangiogenic
CC agents, for the treatment process and diseases involving angiogenesis.
CC Such diseases include cancers such as solid tumours, blood born tumours
CC such as leukemias, tumours metastasis, benign tumours, such as
CC hemangiomas, acoustic acouromas, neurofibromas, trachomas and pyogenic
CC granulomas, rheumatoid arthritis, ocular angiogenic diseases such as
CC diabetic retinopathy, retinopathy of prematurity, macular degeneration,
CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasia,
CC rubrosis, Scler-Webber syndrome, myocardial angiogenesis, plaque
CC neovascularization, telangiectasia, hemophilic joints, angiofibroma,
CC and wound granulation. The fragments are also useful in treatment of
CC disease of excessive or abnormal stimulation of endothelial cells. These
CC diseases include intestinal adhesions, atherosclerosis, scleroderma, and
CC hypertrophic scars. The fragments can be used as birth control agents by
CC preventing vascularization required for embryo transplantation. The
CC present sequence represents the kringle 4 sequence of human plasminogen
CC (angiotensin).
XX
SQ Sequence 84 AA;
Query Match 82.28; Score 430; DB 21; Length 84;
Best Local Similarity 83.38; Pred. No. 1.5e-36;
Matches 70; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
QY 1 VRQYHGNGQSYRGTFSTVTGTCQSSMTPHRHQTPEPNDGLTMNCRNPDADT 60
DB 1 VQDYHGNGQSYRGTFSTVTGTCQSSMTPHRHQTPEPNDGLTMNCRNPDADK 60
QY 61 GPWCFTTDPDSIRWEYCNLTRCSDT 84
DB 61 GPWCFTTDPDSIRWEYCNLTRCSDT 84
RESULT 3
RAP93231
ID AAP93231 standard; peptide: 82 AA.
AC AAP93231;
XX
XX 25-MAR-2003 (updated)
DT 03-APR-1990 (first entry)
XX
XX Plasminogen kringle 4 domain (residues 354-435).
XX
KW Plasminogen; activator; t-PA; fibrin; kringle domain;
KW SCU-PA;
XX
XX Homo sapiens.
XX
XX W08910401-A.
XX
XX 02-NOV-1989.
XX
XX 23-MAR-1989; 89WO-US01255.
XX
XX 22-APR-1988; 88US-0184823.
XX
XX (COLB ) COLLABORATIVE RES INC.
XX
XX Mao JL, Abercrombie DM;
XX
XX WPI; 1989-339965/46.
XX
XX N-PSDB; AAN92238.
XX
XX Modified plasminogen activator - having greater fibrin
PT selectivity and circulating half-life.

```

XX Disclosure; fig 3b; 80pp; English.

PS The sequence is encoded by a synthetic oligonucleotide and is the kringle

CC 4 domain of plasminogen. It is used to construct a modified plasminogen

CC activator (mPA) of greater fibrin selectivity than normal PA. The mPA is

CC used in the treatment of pulmonary embolism, thrombosis, myocardial

CC infarct and strokes.

CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 82 AA;

SQ

Query Match 81.6%; Score 427; DB 10; Length 82;

Best Local Similarity 85.2%; Pred. No. 3.1e-36;

Matches 69; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 VROCYNHGQSYRGFTSTVTGTCQSWSMTPHRRQTPENYPNDGLTMNYCRNPDADT 60

DB 1 VROCYNHGQSYRGFTSTVTGTCQSWSMTPHRRQTPENYPNDGLTMNYCRNPDADT 60

QY 61 GPWCFTTDPSSIRWEYCNLTRC 81

DB 62 GPWCFTTDPSSIRWEYCNLTRC 82

RESULT 4

AA08414

ID AA08414 standard; Protein: 83 AA.

XX

AC AA08414;

DT 20-DEC-2000 (first entry)

XX

DE Amino acid sequence of kringle 4 of human plasminogen.

XX

KW Angiogenesis inhibitor; KED; kringle protein; angiogenesis; tPA;

KW tissue plasminogen activator; tumour; atherosclerosis; arthritis;

KW retinopathy; bronchial vascular congestion; inflammatory bowel disease;

KW adult respiratory distress syndrome; Castellan's disease; psoriasis;

KW hepatitis; aneurysm; renal disease; haemangioma.

OS Homo sapiens.

XX

PN WO200049871-A1.

XX

PD 31-AUG-2000.

XX

PF 24-FEB-2000; 2000WO-US04798.

XX

PR 24-FEB-1999; 99US-0121341.

PR 25-FEB-1999; 99US-0121633.

PR 18-NOV-1999; 99US-0166176.

XX

PA (FORD-) FORD HEALTH SYSTEM HENRY.

XX

PI Dou D, Chopp M, Wang L, Mikkelsen T;

XX

DR WPI; 2000-572016/53.

XX

PT Use of kringle protein and kringle derived from plasminogen and

PT composition comprising kringle proteins for treating tumor and

PT atherosclerosis, arthritis and retinopathy.

XX

PS Disclosure; Fig 6; 163pp; English.

XX

CC The specification describes a human polypeptide which is a potent

CC angiogenesis inhibitor, and is designated KED. KED polypeptide, kringle

CC proteins, or a kringle derived from human tissue plasminogen activator

CC (tPA) protein are used to inhibit angiogenesis. Kringle protein, KED

CC or tPA are useful for treating tumours, as well as atherosclerosis,

CC arthritis, retinopathy and other similar diseases. KED is also useful

CC for the treatment of diseases such as bronchial vascular congestion,

CC inflammatory bowel disease, adult respiratory distress syndrome,

CC Castellan's disease, psoriasis, hepatitis, aneurysm, renal disease

CC and haemangioma. The present sequence represents kringle 4 of human

CC plasminogen, which is used in the course of the invention.

XX Sequence 83 AA;

SQ

Query Match 80.5%; Score 421; DB 21; Length 83;

Best Local Similarity 84.1%; Pred. No. 1.3e-35;

Matches 69; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 VROCYNHGQSYRGFTSTVTGTCQSWSMTPHRRQTPENYPNDGLTMNYCRNPDADT 60

DB 1 VROCYNHGQSYRGFTSTVTGTCQSWSMTPHRRQTPENYPNDGLTMNYCRNPDADT 60

QY 61 GPWCFTTDPSSIRWEYCNLTRC 82

DB 61 GPWCFTTDPSSIRWEYCNLTRC 82

RESULT 5

AA07560

ID AA07560 standard; protein; 78 AA.

XX

AC AA07560;

DT 22-JUN-1997 (first entry)

XX

DE Human kringle 4.

XX

KW angiotatin; plasminogen; kringle; angiogenesis; cancer; arthritis;

KW macular degeneration; diabetic retinopathy.

XX

OS Homo sapiens.

XX

PN WO9635774-A2.

XX

PD 14-NOV-1996.

XX

PF 26-APR-1996; 96WO-US058556.

XX

PR 08-MAR-1996; 96US-0612788.

PR 26-APR-1995; 95US-0429743.

PR 22-FEB-1996; 96US-0605598.

XX

PA (CHIL-) CHILDRENS MEDICAL CENT.

XX

PI Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;

XX

DR WPI; 1996-518662/51.

XX

PT Use of angiotatin fragments or aggregates - for inhibiting

PT endothelial cell proliferation and treating angiogenesis-mediated

PT diseases, e.g. cancer, arthritis or diabetic retinopathy

XX

PS Claim 4; Page 116; 203pp; English.

XX

CC The invention relates to new methods and compositions for

CC inhibiting endothelial cell proliferation, using as active component

CC an angiotatin fragment, a combination of angiotatin fragments, or

CC aggregate angiotatin. The fragment is preferably derived from murine,

CC human, Rhesus, porcine or bovine plasminogen and is a kringle 1,

CC kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle

CC 1-4 or kringle 1-4BKLs protein. The aggregate angiotatin has a Mol. Wt.

CC of 45-65 kD and is derived from a plasminogen fragment beginning at

CC approximately amino acid number 98 of murine, human, Rhesus, porcine or

CC bovine plasminogen. The active component can be used for treating

CC angiogenesis-mediated diseases such as cancer, arthritis, macular

CC degeneration and diabetic retinopathy. It can also be used to develop

CC antibodies for use in diagnosis, detection and therapy.

CC The present sequence, human kringle 4, is a specific angiotatin

CC fragment which can be used in the invention.

XX

SQ Sequence 78 AA;

Query Match 79.78; Score 417; DB 17; Length 78;
 Best Local Similarity 85.98; Pred. No. 3e-35; 6; Indels 0; Gaps 0;
 Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGTCQSWSSMTPHRHQTPTENYDGLTMYCRNPDA7GPW 63
 Db 1 CYHNGQSYRGFTSTVTGTCQSWSSMTPHRHQTPTENYDGLTMYCRNPDA7GPW 60

QY 64 CFTTDPSPRWYCNLRK 81
 Db 61 CFTTDPSPRWYCNLRK 78

RESULT 6
 AAW07559
 ID AAW07559 standard; protein; 78 AA.
 AC AAW07559;
 XX
 XX
 DT 22-JUN-1997 (first entry)
 DE Murine kringle 4.
 XX
 XX
 KW angiotensin; plasminogen; kringle; angiogenesis; cancer; arthritis;
 KW macular degeneration; diabetic retinopathy.
 XX
 OS Mus musculus.
 XX
 PN W09635774-R2.
 XX
 PD 14-NOV-1996.
 XX
 PF 26-APR-1996; 96WO-US05856.
 XX
 PR 08-MAR-1996; 96US-0612788.
 PR 26-APR-1995; 95US-0429743.
 PR 22-FEB-1996; 96US-0605598.
 XX
 XX (CHIL-) CHILDRENS MEDICAL CENT.
 XX
 PI Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;
 XX WPT; 1996-518662/51.
 DR
 XX
 PT Use of angiotensin fragments or aggregates - for inhibiting
 PT endothelial cell proliferation and treating angiogenesis-mediated
 PT diseases, e.g. cancer, arthritis or diabetic retinopathy
 XX
 PS Claim 4; Page 115; 203pp; English.
 XX
 CC The invention relates to new methods and compositions for
 CC inhibiting endothelial cell proliferation, using as active component
 CC an angiotensin fragment, a combination of angiotensin fragments, or
 CC aggregate angiotensin. The fragment is preferably derived from murine,
 CC human, rhesus, porcine or bovine plasminogen and is a kringle 1,
 CC kringle 2, kringle 3, kringle 1-3, kringle 1-2, kringle
 CC 1-4 or kringle 1-4BXLs protein. The aggregate angiotensin has a Mol. wt.
 CC of 45-65 kD and is derived from a plasminogen fragment beginning at
 CC approximately amino acid number 98 of murine, human, rhesus, porcine or
 CC bovine plasminogen. The active component can be used for treating
 CC angiogenesis-mediated diseases such as cancer, arthritis, macular
 CC degeneration and diabetic retinopathy. It can also be used to develop
 CC antibodies for use in diagnosis, detection and therapy.
 CC The present sequence, murine kringle 4, is a specific angiotensin
 CC fragment which can be used in the invention, and represents amino
 CC acids 377-454 of murine plasminogen.
 XX
 XX Sequence 78 AA;

Query Match 70.68; Score 369; DB 17; Length 78;
 Best Local Similarity 73.18; Pred. No. 2.4e-30;
 Matches 57; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGTCQSWSSMTPHRHQTPTENYDGLTMYCRNPDA7GPW 63
 Db 1 CYHNGQSYRGFTSTVTGTCQSWSSMTPHRHQTPTENYDGLTMYCRNPDA7GPW 60

QY 64 CFTTDPSPRWYCNLRK 81
 Db 61 CFTTDPSPRWYCNLRK 78

RESULT 7
 ABP02801
 ID ABP02801 standard; Protein; 86 AA.
 XX
 AC ABP02801;
 XX
 DT 24-JUN-2002 (first entry)
 DE Human ORFX protein sequence SEQ ID NO:5584.
 XX
 XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.
 XX
 OS Homo sapiens.
 XX
 PN W0200192523-A2.
 XX
 PD 06-DEC-2001.
 XX
 XX 29-MAY-2001; 2001WO-US10836.
 XX
 XX 30-MAY-2000; 2000US-206132P.
 PR 29-AUG-2000; 2000US-228716P.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 XX Shimkets RA, Leach MD;
 PI
 XX WPI; 2002-106308/14.
 DR N-PSDB; ABN18553.
 XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders -
 XX
 PS Disclosure; SEQ ID 5584; 1037pp; English.

The present invention describes substantially purified human proteins
 (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 in the specification)). ABN15762 to ABN27252 encode the human ORFX
 proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 treating or preventing a pathology associated with an ORFX-associated
 disorder in humans, and in the manufacture of a medicament for treating a
 syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 sequences can be used in gene therapy. ORFX sequences can be used in the
 treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 osteoarthritis, neurodegenerative disorders, disorders related to organ
 transplantation, cardiovascular diseases, diabetes mellitus, systemic
 lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 storage disease, various immune deficiencies and disorders, infectious
 diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 disease and autoimmune inflammatory eye disease. ORFX proteins are also
 useful for treating burns, incisions, ulcers, for treating osteoporosis,
 bone degenerative disorders, or periodontal disease, and for gut
 protection or regeneration and treatment of lung or liver fibrosis,

The invention relates to new methods and compositions for inhibiting endothelial cell proliferation, using as active component an angiotensin fragment, a combination of angiotensin fragments, or aggregate angiotensin. The fragment is preferably derived from murine, human, Rhesus, porcine or bovine plasminogen and is a kringle 1, kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle 1-4 or kringle 1-4&XLS protein. The aggregate angiotensin has a Mol. Wt. of 45-65 kD and is derived from a plasminogen fragment beginning at approximately amino acid number 98 of murine, human, Rhesus, porcine or bovine plasminogen. The active component can be used for treating angiogenesis-mediated diseases such as cancer, arthritis, macular degeneration and diabetic retinopathy. It can also be used to develop antibodies for use in diagnosis, detection and therapy. The present sequence, murine kringle 3, is a specific angiotensin fragment which can be used in the invention, and represents amino


```
OY 64 CFTTDPESIRWEYCNLTRC 81
Db 61 CTTTSEVKEWETCTIPSC 78

RESULT 14
AAW07553
ID AAW07553 standard; protein; 78 AA.
XX
AC AAW07553;
XX
DT 21-JUN-1997 (first entry)
XX
DE Bovine kringle 2.
XX
KW angiotensin; plasminogen; kringle; angiogenesis; cancer; arthritis;
KW macular degeneration; diabetic retinopathy.
XX
OS Bos taurus.
XX
PN W09635774-A2.
XX
PD 14-NOV-1996.
XX
PF 26-APR-1996; 96WO-US05856.
XX
PR 08-MAR-1996; 96US-0612788.
XX
PR 26-APR-1995; 95US-0429743.
XX
PR 22-FEB-1996; 96US-0605598.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;
XX
XX WPT; 1996-518662/51.
XX
XX Use of angiotensin fragments or aggregates - for inhibiting
PT endothelial cell proliferation and treating angiogenesis-mediated
PT diseases, e.g. cancer, arthritis or diabetic retinopathy
XX
PS Claim 4; Page 110; 203pp; English.
XX
XX The invention relates to new methods and compositions for
CC inhibiting endothelial cell proliferation, using as active component
CC an angiotensin fragment, a combination of angiotensin fragments, or
CC aggregate angiotensin. The fragment is preferably derived from murine,
CC human, Rhesus, porcine or bovine plasminogen and is a kringle 1,
CC kringle 2, kringle 3, kringle 1-3, kringle 1-2, kringle
CC 1-4 or kringle 1-4BKLS protein. The aggregate angiotensin has a Mol. Wt.
CC of 45-65 kD and is derived from a plasminogen fragment beginning at
CC approximately amino acid number 98 of murine, human, Rhesus, porcine or
CC bovine plasminogen. The active component can be used for treating
CC angiogenesis-mediated diseases such as cancer, arthritis, macular
CC degeneration and diabetic retinopathy. It can also be used to develop
CC antibodies for use in diagnosis, detection and therapy.
CC The present sequence, bovine kringle 2, is a specific angiotensin
CC fragment which can be used in the invention, and represents amino
CC acids 88-165 of bovine angiotensin.
XX
SQ Sequence 78 AA;
Query Match 53.0%; Score 277; DB 17; Length 78;
Best Local Similarity 52.6%; Pred. No. 6.1e-21;
Matches 41; Conservative 13; Mismatches 24; Indels 0; Gaps 0;

OY 4 CYHNGQSYRGFTSTVIGRTCSQSWSMTPHRTQPTENYDGLTMVYCRNPADTGP 63
Db 1 CMFCSGENYEGKIATSGRDCQAWDSQSFHAGTIPSKFPNKNKMYCRNPDPGEPRW 60
OY 64 CFTTDPESIRWEYCNLTRC 81
Db 61 CFTTDPQKWEFCDIPTC 78

Search completed: September 29, 2003, 08:33:33
Job time : 83 secs
```

```
RESULT 15
AAB08412
ID AAB08412 standard; Protein; 83 AA.
XX
AC AAB08412;
XX
DT 20-DEC-2000 (first entry)
XX
DE Amino acid sequence of kringle 3 of human plasminogen.
XX
KW Angiogenesis inhibitor; KED; kringle protein; angiogenesis; tPA;
KW tissue plasminogen activator; tumour; atherosclerosis; arthritis;
KW retinopathy; bronchial vascular congestion; inflammatory bowel disease;
KW adult respiratory distress syndrome; Castleman's disease; psoriasis;
KW hepatitis; aneurysm; renal disease; haemangioma.
XX
OS Homo sapiens.
XX
PN W0200049871-A1.
XX
PD 31-AUG-2000.
XX
PF 24-FEB-2000; 2000WO-US04798.
XX
PR 24-FEB-1999; 99US-0121341.
XX
PR 25-FEB-1999; 99US-0121633.
XX
PR 18-NOV-1999; 99US-0166176.
XX
PA (FORD-) FORD HEALTH SYSTEM HENRY.
XX
PI Dou D, Chopp M, Wang L, Mikkelsen T;
XX
XX WPT; 2000-572016/53.
XX
XX Use of kringle protein and kringle derived from plasminogen and
PT composition comprising kringle proteins for treating tumor and
PT atherosclerosis, arthritis and retinopathy
XX
PS Disclosure; Fig 6; 163pp; English.
XX
XX The specification describes a human polypeptide which is a potent
CC angiogenesis inhibitor, and is designated KED. KED polypeptide, kringle
CC proteins, or a kringle derived from human tissue plasminogen activator
CC (tPA) protein are used to inhibit angiogenesis. Kringle protein, KED
CC or tPA are useful for treating tumours, as well as atherosclerosis,
CC arthritis, retinopathy and other similar diseases. KED is also useful
CC for the treatment of diseases such as bronchial vascular congestion,
CC inflammatory bowel disease, adult respiratory distress syndrome,
CC Castleman's disease, psoriasis, hepatitis, aneurysm, renal disease
CC and haemangioma. The present sequence represents kringle 3 of human
CC plasminogen, which is used in the course of the invention.
XX
SQ Sequence 83 AA;
Query Match 52.8%; Score 276; DB 21; Length 83;
Best Local Similarity 55.0%; Pred. No. 8.3e-21;
Matches 44; Conservative 8; Mismatches 28; Indels 0; Gaps 0;

OY 3 QCYHNGQSYRGFTSTVIGRTCSQSWSMTPHRTQPTENYDGLTMVYCRNPADTGP 62
Db 3 QCLKPGGENYRGNNNAVIVSGTSCQHSAGTPTHTNRTPEFPCCKNLNDENYCRNPDKRAP 62
OY 63 WCFTTDPESIRWEYCNLTRCS 82
Db 63 WCHTINSQVRWEYCKTPSCS 82

Search completed: September 29, 2003, 08:33:33
Job time : 83 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:32:08 ; Search time 28 Seconds
(without alignments)
134.488 Million cell updates

Title: US-10-088-548-6
 Perfect score: 523
 Sequence: 1 VRQCYHGNGQSVRGTFSTV.....SIRWEYCNLTRCSDEGTGVV 89

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 223717

```
Minimum DB seq length: 0
Maximum DB seq length: 89
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database : Issued_patents_Aa:*
1: /cgn2_6/ptodata1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata1/1aa/FCTUS_COMB.pep.*
6: /cgn2_6/ptodata1/1aa/backfiles1.pep.*
```

SUMMARIES

Result No.	Query			ID	Description
	Score	Match	Length		
1	430	82.2	84	US-09-348-953-23	Sequence 2, Appl
2	417	79.7	78	US-08-612-788-23	Sequence 23, Appl
3	417	79.7	78	US-08-763-528A-5	Sequence 5, Appl
4	417	79.7	78	US-09-066-028-23	Sequence 23, Appl
5	417	79.7	78	US-09-335-325-23	Sequence 23, Appl
6	369	70.6	78	US-08-612-788-22	Sequence 22, Appl
7	369	70.6	78	US-09-066-028-22	Sequence 22, Appl
8	369	70.6	78	US-09-335-325-22	Sequence 22, Appl
9	313	59.8	78	US-08-612-788-17	Sequence 17, Appl
10	313	59.8	78	US-09-066-028-17	Sequence 17, Appl
11	313	59.8	78	US-09-335-325-17	Sequence 17, Appl
12	299	57.2	78	US-08-612-788-20	Sequence 20, Appl
13	299	57.2	78	US-09-066-028-20	Sequence 20, Appl
14	299	57.2	78	US-09-335-325-20	Sequence 20, Appl
15	285	54.5	78	US-08-612-788-18	Sequence 18, Appl
16	285	54.5	78	US-09-066-028-18	Sequence 18, Appl
17	285	54.5	78	US-09-335-325-18	Sequence 18, Appl
18	283	54.1	78	US-08-612-788-19	Sequence 19, Appl
19	283	54.1	78	US-09-066-028-19	Sequence 19, Appl
20	283	54.1	78	US-09-335-325-19	Sequence 19, Appl
21	281	53.7	78	US-08-612-788-21	Sequence 21, Appl
22	281	53.7	78	US-08-763-528A-4	Sequence 4, Appl
23	281	53.7	78	US-09-066-028-21	Sequence 21, Appl
24	281	53.7	78	US-09-335-325-21	Sequence 21, Appl
25	277	53.0	78	US-08-612-788-16	Sequence 16, Appl
26	277	53.0	78	US-09-066-028-16	Sequence 16, Appl
27	277	53.0	78	US-09-335-325-16	Sequence 16, Appl

ALIGNMENTS

```

RESULT 1
US-09-348-953-2
  / Sequence 2, Application US/09348953
  / Patent No. 6538103
  / GENERAL INFORMATION:
  / APPLICANT: Ji, Richard Wei-Dong
  / APPLICANT: Trail, Pamela A.
  / TITLE OF INVENTION: LYSINE BINDING FRAGMENT
  / FILE REFERENCE: DB1sequences
  / CURRENT APPLICATION NUMBER: US/09/348,953
  / CURRENT FILING DATE: 1999-07-07
  / PRIOR APPLICATION NUMBER: 60/092,831
  / PRIOR FILING DATE: 1998-07-14
  / NUMBER OF SEQ ID NOS: 8
  / SOFTWARE: Patentin Ver. 2.0
  / SEQ ID NO 2
  / LENGTH: 84
  / TYPE: PRT
  / ORGANISM: human
  / US-09-348-953-2

```

[illegible]

```
QY      61 GPWCFTDPSIRWEYCNLTRCSDT 84
      |||||:|||||:|
pb      61 GPWCFTDPSVRWEYCNLKKCSGT 84
```

```

RESULT 2
US-08-612-788-23
: Sequence 23, Application US/08612788
: Patent No. 5837682
: GENERAL INFORMATION:
: APPLICANT: Folkman, M. Judah
: APPLICANT: O'Reilly, Micheal
: APPLICANT: Cao, Yihai
: APPLICANT: Sim, B. Kim Lee
: TITLE OF INVENTION: Angiostatin Fragments and Method of Use
: NUMBER OF SEQUENCES: 45
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Jones & Askew
: STREET: 191 Peachtree Street, 37th Floor
: CITY: Atlanta
: STATE: Georgia

```

```
/
/ COUNTRY: U.S.
/ ZIP: 30303-1769
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE:
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 23:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 78 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ IMMEDIATE SOURCE:
/ CLONE: K4
/
US-08-612-788-23

Query Match 79.7%; Score 417; DB 2; Length 78;
Best Local Similarity 85.9%; Pred. No. 3.1e-40;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CYHNGQSYRGTSFTVTGTCQSWSSMTPHRRHQTENYPNDGLTMNYCRNPDADTGPW 63
Db 1 CYHGDGQSYRGTSFTVTGTCQSWSSMTPHRRHQTENYPNAGLTMYCRNPDADKGPW 60

Qy 64 CFTTDPISRWEYCNLRK 81
Db 61 CFTTDPISRWEYCNLRK 78

RESULT 3
US-08-763-528A-5
; Sequence 5, Application US/08763528A
; Patent No. 5854221
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Endothelial Cell Proliferation Inhibitor
; TITLE OF INVENTION: and Method of Use
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew, LLP
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/763,528A
; FILING DATE: 12-DEC-1996
; CLASSIFICATION: 530
```

```
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05940-0251
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 5:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 78 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: internal
/ FEATURE:
/ NAME/KEY: Protein
/ LOCATION: 1..78
/ OTHER INFORMATION: /note= "Kringle 4 - Figure 3"
/
US-08-763-528A-5

Query Match 79.7%; Score 417; DB 2; Length 78;
Best Local Similarity 85.9%; Pred. No. 3.1e-40;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CYHNGQSYRGTSFTVTGTCQSWSSMTPHRRHQTENYPNDGLTMNYCRNPDADTGPW 63
Db 1 CYHGDGQSYRGTSFTVTGTCQSWSSMTPHRRHQTENYPNAGLTMYCRNPDADKGPW 60

Qy 64 CFTTDPISRWEYCNLRK 81
Db 61 CFTTDPISRWEYCNLRK 78

RESULT 4
US-09-066-028-23
; Sequence 23, Application US/09066028
; Patent No. 6024588
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
```

```

; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: K4
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-066-028-23

Query Match 79.7%; Score 417; DB 3; Length 78;
Best Local Similarity 85.9%; Pred. No. 3.le-40;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGTFSTVIGTCQSWSSMTPHRHQRTPENYPNDGLTMNYCNPADADGPW 63
Db 1 CYHGDGQSYRGTSSTTTGKCKQSWSSMTPHRHQRTPENYPNAGLTMYCNPADADGPW 60
QY 64 CFTTDPSPRWYCNLTKC 81
Db 61 CFTTDPSPRWYCNLTKC 78

RESULT 5
US-09-335-325-23
; Sequence 23, Application US/09335325
; Patent No. 6521439
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; O'Reilly, Micheal
; Cao, Yihai
; Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K4
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-08-612-788-22

; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; IMMEDIATE SOURCE:
; CLONE: K4
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-335-325-23

Query Match 79.7%; Score 417; DB 4; Length 78;
Best Local Similarity 85.9%; Pred. No. 3.le-40;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGTFSTVIGTCQSWSSMTPHRHQRTPENYPNDGLTMNYCNPADADGPW 63
Db 1 CYHGDGQSYRGTSSTTTGKCKQSWSSMTPHRHQRTPENYPNAGLTMYCNPADADGPW 60
QY 64 CFTTDPSPRWYCNLTKC 81
Db 61 CFTTDPSPRWYCNLTKC 78

RESULT 6
US-08-612-788-22
; Sequence 22, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; O'Reilly, Micheal
; Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K4
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-08-612-788-22
```

```
Query Match 70.6%; Score 369; DB 2; Length 78;
Best Local Similarity 73.1%; Pred. No. 8.6e-35;
Matches 57; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTFTVTGTCQSWSSMTPHRRHQRTPENYDGLTMNYCRNPDADTGPW 63
  || :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 1 CYQSDGQSYRGFTSTTTTGKQSWAAMPFPHRSKTPENFDAGLENNYCRNPDGDKGPW 60

QY 64 CFTTDPSSIRWEYCNLTRC 81
Db 61 CTTTDFSVRWEYCNLKRK 78

RESULT 7
US-09-066-028-22
; Sequence 22, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yibai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K4
US-09-066-028-22

Query Match 70.6%; Score 369; DB 3; Length 78;
Best Local Similarity 73.1%; Pred. No. 8.6e-35;
Matches 57; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTFTVTGTCQSWSSMTPHRRHQRTPENYDGLTMNYCRNPDADTGPW 63
  || :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 1 CYQSDGQSYRGFTSTTTTGKQSWAAMPFPHRSKTPENFDAGLENNYCRNPDGDKGPW 60

QY 64 CFTTDPSSIRWEYCNLTRC 81
Db 61 CTTTDFSVRWEYCNLKRK 78

RESULT 8
US-09-335-325-22
; Sequence 22, Application US/09335325
; Patent No. 6521439
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yibai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS: <unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K4
US-09-335-325-22

Query Match 70.6%; Score 369; DB 4; Length 78;
Best Local Similarity 73.1%; Pred. No. 8.6e-35;
Matches 57; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTFTVTGTCQSWSSMTPHRRHQRTPENYDGLTMNYCRNPDADTGPW 63
  || :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 1 CYQSDGQSYRGFTSTTTTGKQSWAAMPFPHRSKTPENFDAGLENNYCRNPDGDKGPW 60

QY 64 CFTTDPSSIRWEYCNLTRC 81
Db 61 CTTTDFSVRWEYCNLKRK 78
```


RESULT 9

```

US-08-612-788-17
; Sequence 17, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiotstatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0136
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K3
; US-08-612-788-17

Query Match 59.8%; Score 313; DB 2; Length 78;
Best Local Similarity 61.5%; Pred.No. 1.9e-28;
Matches 48; Conservative 9; Mismatches 21; Indels

QY 4 CYHGNGSYRGTEFTVTGTCQSSMTPHRHORTPENYPNDGLTMYNCR
| | :||| | :||| | :||| | :||| | :||| | :||| | :|||
Db 1 CLKRGENYSRGTVTSVSKTCQRSEQTPHHNRTPFNFCNMLENCVCRC
| :||| | :||| | :||| | :||| | :||| | :||| | :|||
QY 64 CFTTDPSIRWEYNLRTC 81
| :||| | :||| | :||| | :||| | :||| | :||| | :|||
Db 61 CYTTDSQLRWEYCEIPSC 78

RESULT 10
US-09-066-028-17
; Sequence 17, Application US/09066028
; Patent No. 5024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiotstatin Fragments and Method of Use

```

NUMBER OF SEQUENCES. 45

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askew
STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: 05213-0126
APPLICATION NUMBER: US/09/066,028
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/612,788
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren, William L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 05213-0126
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 78 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Murine
IMMEDIATE SOURCE:
CLONE: K3
US-09-066-028-17

Query Match 59.8%; Score 313; DB 3; Length 78;
Best Local Similarity 61.5%; Pred. NO. 1.9e-28;
Matches 46; Conservative 9; Mismatches 21; Indels 0; Gaps 0;

QY 4 CYHGNQSYRGTFSTVTGRTCSWSMTPHRRHQRTPENYPNDGLTMYCRNPADTGF 63
| | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CLKRGENTRGTVTVTSVKTKQWSEQIPHRHNTPTNFPCKNLEENYCRNPDGSTAPW 60
| | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 64 CFTDFSIIRWYCNLTRC 81
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CYTDSQLRWEYCEIPSC 78

RESULT 11
US-09-335-325-17
; Sequence 17, Application US/09335325
; Patent No. 6521439
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; O'Reilly, Micheal
; Cao, Yihai
; Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.

```

```
;
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: K3
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-335-325-17

Query Match 59.8%; Score 313; DB 4; Length 78;
Best Local Similarity 61.5%; Pred. No. 1.9e-28;
Matches 48; Conservative 9; Mismatches 21; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGTCQSWSPHQRHPTPENYDGLTMYCRNPDADTGPW 63
DB 1 CLKGRGENYRGVTSVSGTQCRWSAQSPKHKHRTPEPCKNLEENYCRNPDGETAPW 60

QY 64 CFTTDPSSIRWEYCNLTRC 81
DB 61 CYTTDSQLRWEYCEIPSC 78

RESULT 12
US-08-612-788-20
; Sequence 20, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
```

```
;
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Porcine
; IMMEDIATE SOURCE:
; CLONE: K3
; US-08-612-788-20

Query Match 57.2%; Score 299; DB 2; Length 78;
Best Local Similarity 56.4%; Pred. No. 7.4e-27;
Matches 44; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGTCQSWSPHQRHPTPENYDGLTMYCRNPDADTGPW 63
DB 1 CLKGRGENYRGVTSVSGTQCRWSAQSPKHKHRTPEPCKNLEENYCRNPDGETAPW 60

QY 64 CFTTDPSSIRWEYCNLTRC 81
DB 61 CYTTDSQLRWEYCEIPSC 78

RESULT 13
US-09-066-028-20
; Sequence 20, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
```

REFERENCE/DOCKET NUMBER: 05213-0126
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 78 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
IMMEDIATE SOURCE:
CLONE: K3

US-09-066-028-20
Query Match 57.2%; Score 299; DB 3; Length 78;
Best Local Similarity 56.4%; Pred. No. 7.4e-27;
Matches 44; Conservative 12; Mismatches 22; Indels 0; Gaps 0;
QY 4 CYHNGQSVRGTFSTVTGRTCCQSSWMTPHRHQRTPEPNYNDGLTMNYCRNPDAITGPW 63
DB 1 CLARGENYRGTVSVTASGHTCQWSAQSPKHNRTPEPNFCKNLEENYCRNPDPGETAPW 60
QY 64 CFTTDFSIWXYCNLTRC 81
DB 61 CYTTDSEVRWDYCKIPSC 78

RESULT 14
US-09-335-325-20
Sequence 20, Application US/09333325
Patent No. 6521439
GENERAL INFORMATION:
APPLICANT: Folkman, M. Judah
O'Reilly, Michael
Cao, Yihai
Sim, B. Kim Lee
TITLE OF INVENTION: Angiostatin Fragments and Method of Use
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askew
STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/335,325
FILING DATE: 17-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/612,788
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Warren, William L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 05213-0126
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 78 amino acids

TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Porcine
IMMEDIATE SOURCE:
CLONE: K3
SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-335-325-20
Query Match 57.2%; Score 299; DB 4; Length 78;
Best Local Similarity 56.4%; Pred. No. 7.4e-27;
Matches 44; Conservative 12; Mismatches 22; Indels 0; Gaps 0;
QY 4 CYHNGQSVRGTFSTVTGRTCCQSSWMTPHRHQRTPEPNYNDGLTMNYCRNPDAITGPW 63
DB 1 CLARGENYRGTVSVTASGHTCQWSAQSPKHNRTPEPNFCKNLEENYCRNPDPGETAPW 60
QY 64 CFTTDFSIWXYCNLTRC 81
DB 61 CYTTDSEVRWDYCKIPSC 78
RESULT 15
US-08-612-788-18
Sequence 18, Application US/08612788
Patent No. 5837682
GENERAL INFORMATION:
APPLICANT: Folkman, M. Judah
O'Reilly, Michael
Cao, Yihai
Sim, B. Kim Lee
TITLE OF INVENTION: Angiostatin Fragments and Method of Use
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askew
STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,788
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Warren, William L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 05213-0126
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 78 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens

```
/ IMMEDIATE SOURCE:
/ CLONE: K3
US-08-612-788-18

Query Match 54.5%; Score 285; DB 2; Length 78;
Best Local Similarity 55.1%; Pred. No. 2.8e-25;
Matches 43; Conservative 10; Mismatches 25; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGTFSTVTYGTCTQSWSSMTPIRPHORTPENYPNDGLTMNYCRNPDADTGPW 63
Db | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
1 CLKGTGENYRGNAVTVSGHTCOHWSAQTPHTNRTPENFPKSLDENYCRNPDGKRAPW 60

QY 64 CFTTDPISIRWEYCNLRC 81
Db | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
61 CHTNSQVRWEYCKIPSC 78
```

Search completed: September 29, 2003, 08:37:06
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:35:54 : Search time 61 Seconds
(without alignments)
220.760 Million cell updates/sec

Title: US-10-088-548-6

Perfect score: 523

Sequence: 1 VRQYHGNGSQVRGTFSTV.....SIRWEYCNLTRCSDIEGTGV 89

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 201696

Minimum DB seq length: 0

Maximum DB seq length: 89

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	417	79.7	78	9	US-09-753-064-5
2	417	79.7	78	9	US-09-761-120-23
3	417	79.7	78	10	US-09-335-325-23
4	417	79.7	78	12	US-10-267-137-9
5	417	79.7	78	15	US-10-131-241-23
6	369	70.6	78	9	US-09-761-120-22
7	369	70.6	78	10	US-09-335-325-22
8	369	70.6	78	15	US-10-131-241-22
9	313	59.8	78	9	US-09-761-120-17
10	313	59.8	78	10	US-09-335-325-17
11	313	59.8	78	15	US-10-131-241-17
12	299	57.2	78	9	US-09-761-120-20
13	299	57.2	78	10	US-09-335-325-20
14	299	57.2	78	15	US-10-131-241-20
15	285	54.5	78	10	US-09-335-325-18

16	285	54.5	78	15	US-10-131-241-18	Sequence 18, Appl
17	283	54.1	78	9	US-09-761-120-19	Sequence 19, Appl
18	283	54.1	78	10	US-09-335-325-19	Sequence 19, Appl
19	283	54.1	78	15	US-10-131-241-19	Sequence 19, Appl
20	281	53.7	78	9	US-09-753-064-4	Sequence 4, Appl
21	281	53.7	78	9	US-09-761-120-18	Sequence 18, Appl
22	281	53.7	78	9	US-09-761-120-21	Sequence 21, Appl
23	281	53.7	78	10	US-09-335-325-21	Sequence 21, Appl
24	281	53.7	78	15	US-10-131-241-21	Sequence 21, Appl
25	277	53.0	78	9	US-03-761-120-16	Sequence 16, Appl
26	277	53.0	78	10	US-09-335-325-16	Sequence 16, Appl
27	277	53.0	78	15	US-10-131-241-16	Sequence 16, Appl
28	275	52.6	78	12	US-10-267-137-12	Sequence 12, Appl
29	274	52.4	78	9	US-09-761-120-15	Sequence 15, Appl
30	274	52.4	78	10	US-09-335-325-15	Sequence 15, Appl
31	274	52.4	78	15	US-10-131-241-15	Sequence 15, Appl
32	268.5	51.3	79	9	US-09-761-120-9	Sequence 9, Appl
33	268.5	51.3	79	10	US-09-335-325-9	Sequence 9, Appl
34	268.5	51.3	79	15	US-10-131-241-9	Sequence 9, Appl
35	267.5	51.1	79	9	US-09-753-064-2	Sequence 2, Appl
36	267.5	51.1	79	9	US-09-761-120-8	Sequence 8, Appl
37	267.5	51.1	79	10	US-09-335-325-8	Sequence 8, Appl
38	267.5	51.1	79	12	US-10-267-137-7	Sequence 7, Appl
39	267.5	51.1	79	15	US-10-131-241-8	Sequence 8, Appl
40	264	50.5	78	10	US-09-335-325-13	Sequence 13, Appl
41	264	50.5	78	15	US-10-131-241-13	Sequence 13, Appl
42	262	50.1	78	9	US-09-761-120-14	Sequence 14, Appl
43	262	50.1	78	10	US-09-335-325-14	Sequence 14, Appl
44	262	50.1	78	15	US-10-131-241-14	Sequence 14, Appl
45	261	49.9	78	9	US-09-753-064-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-753-064-5
; Sequence 5, Application US/09753064
; Patent No. US20010016644A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; Folkman, M. Judah
; TITLE OF INVENTION: Endothelial Cell Proliferation Inhibitor
and Method of Use
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew, LLP
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/753,064
; FILING DATE: 29-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/763,528
; FILING DATE: 12-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05940-0251
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:

```
/
/ LENGTH: 78 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: internal
/ FEATURE:
/ NAME/KEY: Protein
/ LOCATION: 1..78
/ OTHER INFORMATION: /note= "Kringle 4 - Figure 3"
/ SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-753-064-5

Query Match          79.7%; Score 417; DB 9; Length 78;
Best Local Similarity 85.9%; Pred. No. 5.6e-39;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGRTCSWSSMTPHRRHQRTPENYPNDGLTMNYCRNPADATGPW 63
   |||||:||||| ||| ||: |||||:|||||:|||||:|||||:|||||:|||||
Db 1 CYHGDGQSYRGFTSTVTGRTCSWSSMTPHRRHQRTPENYPNDGLTMNYCRNPADATGPW 60
   |||||:||||| ||| ||: |||||:|||||:|||||:|||||:|||||:|||||

QY 64 CFTTDPSSIRWEYCNLTRC 81
   |||||:||||| ||| ||: |||||:|||||:|||||:|||||:|||||
Db 61 CFTTDPSSIRWEYCNLKKC 78
   |||||:||||| ||| ||: |||||:|||||:|||||:|||||:|||||

RESULT 2
US-09-761-120-23
; Sequence 23, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasminogen
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 23
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Kringle 4
US-09-761-120-23

Query Match          79.7%; Score 417; DB 9; Length 78;
Best Local Similarity 85.9%; Pred. No. 5.6e-39;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGRTCSWSSMTPHRRHQRTPENYPNDGLTMNYCRNPADATGPW 63
   |||||:||||| ||| ||: |||||:|||||:|||||:|||||:|||||
Db 1 CYHGDGQSYRGFTSTVTGRTCSWSSMTPHRRHQRTPENYPNDGLTMNYCRNPADATGPW 60
   |||||:||||| ||| ||: |||||:|||||:|||||:|||||:|||||

QY 64 CFTTDPSSIRWEYCNLTRC 81
   |||||:||||| ||| ||: |||||:|||||:|||||:|||||
Db 61 CFTTDPSSIRWEYCNLKKC 78
   |||||:||||| ||| ||: |||||:|||||:|||||:|||||:|||||

RESULT 3
US-09-335-325-23
; Sequence 23, Application US/09335325
; Patent No. US2002016471A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
```

```
/
/ O'Reilly, Micheal
/ Cao, Yihai
/ Sim, B. Kim Lee
/ TITLE OF INVENTION: Argicstatin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S.
/ ZIP: 30303-1769
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/335,325
/ FILING DATE: 17-Jun-1999
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 23:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 78 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ IMMEDIATE SOURCE:
/ CLONE: K4
/ SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-335-325-23

Query Match          79.7%; Score 417; DB 10; Length 78;
Best Local Similarity 85.9%; Pred. No. 5.6e-39;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGFTSTVTGRTCSWSSMTPHRRHQRTPENYPNDGLTMNYCRNPADATGPW 63
   |||||:||||| ||| ||: |||||:|||||:|||||:|||||:|||||
Db 1 CYHGDGQSYRGFTSTVTGRTCSWSSMTPHRRHQRTPENYPNDGLTMNYCRNPADATGPW 60
   |||||:||||| ||| ||: |||||:|||||:|||||:|||||:|||||

QY 64 CFTTDPSSIRWEYCNLTRC 81
   |||||:||||| ||| ||: |||||:|||||:|||||:|||||
Db 61 CFTTDPSSIRWEYCNLKKC 78
   |||||:||||| ||| ||: |||||:|||||:|||||:|||||:|||||

RESULT 4
US-10-267-137-9
; Sequence 9, Application US/10267137
; Publication No. US20030148950A1
; GENERAL INFORMATION:
; APPLICANT: Xin, Li
; APPLICANT: Li, Zai-Ping
; APPLICANT: Gan, Ren-bao
; APPLICANT: Zhou, Qing-wei
; APPLICANT: Xu, Ren
; TITLE OF INVENTION: KRINGLE DOMAIN 1 OF HUMAN HEPATOCYTE
; TITLE OF INVENTION: GROWTH FACTOR AND USES THEREFOR
```

```
; FILE REFERENCE: 524282000100
; CURRENT APPLICATION NUMBER: US/10/267,137
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 60/328,329
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence alignment of kringle domains of
; OTHER INFORMATION: plasminogen and HGF
US-10-267-137-9

Query Match          79.7%; Score 417; DB 12; Length 78;
Best Local Similarity 85.9%; Pred. No. 5.6e-39;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGCTFTTGTCTCQSSMTPHRHQRTPEPNYNDGLTMNYCRNPADTGPW 63
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 CHHGQSYRGCTFTTGTCTCQSSMTPHRHQRTPEPNYNDGLTMNYCRNPADTGPW 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 64 CFTTDPsirweyCNLRC 81
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 CFTTDPsirweyCNLKKC 78
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
US-10-131-241-23
; Sequence 23, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-23

Query Match          79.7%; Score 417; DB 15; Length 78;
Best Local Similarity 85.9%; Pred. No. 5.6e-39;
Matches 67; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGCTFTTGTCTCQSSMTPHRHQRTPEPNYNDGLTMNYCRNPADTGPW 63
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 CHHGQSYRGCTFTTGTCTCQSSMTPHRHQRTPEPNYNDGLTMNYCRNPADTGPW 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 64 CFTTDPsirweyCNLRC 81
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 CFTTDPsirweyCNLKKC 78
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
US-09-761-120-22
; Sequence 22, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmi
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Kringle 4
US-09-761-120-22

Query Match          70.8%; Score 369; DB 9; Length 78;
Best Local Similarity 73.1%; Pred. No. 1.2e-33;
Matches 57; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGCTFTTGTCTCQSSMTPHRHQRTPEPNYNDGLTMNYCRNPADTGPW 63
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 CHHGQSYRGCTFTTGTCTCQSSMTPHRHQRTPEPNYNDGLTMNYCRNPADTGPW 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 64 CFTTDPsirweyCNLRC 81
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 CFTTDPsirweyCNLKKC 78
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 7
US-09-335-325-22
; Sequence 22, Application US/09335325
; Patent No. US20020164717A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
```

```
/
/   LENGTH: 78 amino acids
/   TYPE: amino acid
/   STRANDEDNESS: <Unknown>
/   TOPOLOGY: linear
/   MOLECULE TYPE: protein
/   HYPOTHEICAL: NO
/   ANTI-SENSE: NO
/   FRAGMENT TYPE: N-terminal
/   ORIGINAL SOURCE:
/     ORGANISM: Murine
/     IMMEDIATE SOURCE:
/     CLONE: K4
/   SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-335-325-22

Query Match      70.6%; Score 369; DB 10; Length 78;
Best Local Similarity 73.1%; Pred. No. 1.2e-33;
Matches 57; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGTFSTVTGRTCSQSSMTPHRHQRTPENYNDGLTMYCRNPDADTGPW 63
Db 1 CYQSDGQSYRGTSSTITGKCKQSWAAMPFHRHSKTPENFPDAGLEMYCRNPDGKGPW 60

QY 64 CFTTDPSSIRWEYCNLTRC 81
Db 61 CYTTDPSVRWEYCNLKR 78

RESULT 8
US-10-131-241-22
/ Sequence 22, Application US/10131241
/ Publication No. US20030012792A1
/ GENERAL INFORMATION:
/ APPLICANT: Holaday, John W.
/ TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
/ FILE REFERENCE: 05213-0344 43170-271565
/ CURRENT APPLICATION NUMBER: US/10/131,241
/ CURRENT FILING DATE: 2002-07-22
/ PRIOR APPLICATION NUMBER: US 09/413,049
/ PRIOR FILING DATE: 1999-10-06
/ PRIOR APPLICATION NUMBER: US 09/316,802
/ PRIOR FILING DATE: 1999-05-21
/ PRIOR APPLICATION NUMBER: US 60/086,586
/ PRIOR FILING DATE: 1998-05-22
/ NUMBER OF SEQ ID NOS: 65
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 22
/ LENGTH: 78
/ TYPE: PRT
/ ORGANISM: Murinae sp.
US-10-131-241-22

Query Match      70.6%; Score 369; DB 15; Length 78;
Best Local Similarity 73.1%; Pred. No. 1.2e-33;
Matches 57; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGTFSTVTGRTCSQSSMTPHRHQRTPENYNDGLTMYCRNPDADTGPW 63
Db 1 CYQSDGQSYRGTSSTITGKCKQSWAAMPFHRHSKTPENFPDAGLEMYCRNPDGKGPW 60

QY 64 CFTTDPSSIRWEYCNLTRC 81
Db 61 CYTTDPSVRWEYCNLKR 78

RESULT 9
US-09-761-120-17
/ Sequence 17, Application US/09761120
/ Patent No. US20020037847A1
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
```

```
/ APPLICANT: O'Reilly, Michael
/ TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmin
/ FILE REFERENCE: 05940-0151 (43171-252068)
/ CURRENT APPLICATION NUMBER: US/09/761,120
/ CURRENT FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: 09/309,821
/ PRIOR FILING DATE: 1999-05-11
/ PRIOR APPLICATION NUMBER: 08/866,735
/ PRIOR FILING DATE: 1997-05-30
/ NUMBER OF SEQ ID NOS: 47
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 17
/ LENGTH: 78
/ TYPE: PRT
/ ORGANISM: Murinae gen. sp.
/ FEATURE:
/ NAME/KEY: misc_feature
/ OTHER INFORMATION: Kringle 3
US-09-761-120-17

Query Match      59.8%; Score 313; DB 9; Length 78;
Best Local Similarity 61.5%; Pred. No. 1.9e-27;
Matches 48; Conservative 9; Mismatches 21; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGTFSTVTGRTCSQSSMTPHRHQRTPENYNDGLTMYCRNPDADTGPW 63
Db 1 CLKGRGENYRGTVSVTSGKTCQWSEQPHRHRTPENPKNLEENYCRNPDGCTAPW 60

QY 64 CFTTDPSSIRWEYCNLTRC 81
Db 61 CYTTDPSQLRWEYCEIPSC 78

RESULT 10
US-09-335-325-17
/ Sequence 17, Application US/09335325
/ Patent No. US20020164717A1
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ O'Reilly, Micheal
/ Cao, Yihai
/ Sim, B. Kim Lee
/ TITLE OF INVENTION: Angiostatin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S.
/ ZIP: 30303-1769
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/335,325
/ FILING DATE: 17-Jun-1999
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 17:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 78 amino acids
```



```
/
/ TYPE: amino acid
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Murine
/ CLONE: K3
/ IMMEDIATE SOURCE:
/ SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-335-325-17

Query Match          59.8%; Score 313; DB 10; Length 78;
Best Local Similarity 61.5%; Pred. No. 1.9e-27;
Matches 48; Conservative 9; Mismatches 21; Indels 0; Gaps 0;

QY 4 CYHNGOSYRGTFSTVTGRTCSWSSMTPHRHQRTPEYNPNDGLTMNYCRNPADTGPW 63
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CLKGGENYRGTVSVTSKTCQWSEQTTPHHRNTPENFPCKNLENYCRNPGETAPW 60

QY 64 CFTTDPSTIRWEYCNLRFC 81
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CYTTDSQLRWEYCEIPSC 78

RESULT 11
US-10-131-241-17
/ Sequence 17, Application US/10131241
/ Publication No. US20030012792A1
/ GENERAL INFORMATION:
/ APPLICANT: Holaday, John W.
/ TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
/ TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
/ FILE REFERENCE: 05213-0344 43170-271565
/ CURRENT APPLICATION NUMBER: US/10/131,241
/ CURRENT FILING DATE: 2002-07-22
/ PRIOR APPLICATION NUMBER: US 09/413,049
/ PRIOR FILING DATE: 1999-10-06
/ PRIOR APPLICATION NUMBER: US 09/316,802
/ PRIOR FILING DATE: 1999-05-21
/ PRIOR APPLICATION NUMBER: US 60/086,586
/ PRIOR FILING DATE: 1998-05-22
/ NUMBER OF SEQ ID NOS: 65
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 17
/ LENGTH: 78
/ TYPE: PRT
/ ORGANISM: Murine
US-10-131-241-17

Query Match          59.8%; Score 313; DB 15; Length 78;
Best Local Similarity 61.5%; Pred. No. 1.9e-27;
Matches 48; Conservative 9; Mismatches 21; Indels 0; Gaps 0;

QY 4 CYHNGOSYRGTFSTVTGRTCSWSSMTPHRHQRTPEYNPNDGLTMNYCRNPADTGPW 63
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CLKGGENYRGTVSVTSKTCQWSEQTTPHHRNTPENFPCKNLENYCRNPGETAPW 60

QY 64 CFTTDPSTIRWEYCNLRFC 81
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CYTTDSQLRWEYCEIPSC 78

RESULT 12
US-09-761-120-20
/ Sequence 20, Application US/09761120
/ Patent No. US2002037847A1
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ APPLICANT: O'Reilly, Michael
```

```
/
/ TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmi
/ FILE REFERENCE: 05940-0151 (43171-252088)
/ CURRENT APPLICATION NUMBER: US/09/761,120
/ CURRENT FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: 09/309,821
/ PRIOR FILING DATE: 1999-05-11
/ PRIOR APPLICATION NUMBER: 08/866,735
/ PRIOR FILING DATE: 1997-05-30
/ NUMBER OF SEQ ID NOS: 47
/ SOFTWARE: Patentin version 3.0
/ SEQ ID NO 20
/ LENGTH: 78
/ TYPE: PRT
/ ORGANISM: Sus sp.
/ FEATURE:
/ NAME/KEY: misc.feature
/ OTHER INFORMATION: Kringle 3
US-09-761-120-20

Query Match          57.2%; Score 299; DB 9; Length 78;
Best Local Similarity 56.4%; Pred. No. 6.9e-26;
Matches 44; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

QY 4 CYHNGOSYRGTFSTVTGRTCSWSSMTPHRHQRTPEYNPNDGLTMNYCRNPADTGPW 63
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CLKGGENYRGTVSVTSKTCQWSEQTTPHHRNTPENFPCKNLENYCRNPGETAPW 60

QY 64 CFTTDPSTIRWEYCNLRFC 81
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CYTTDSVLRWEYCEIPSC 78

RESULT 13
US-09-335-325-20
/ Sequence 20, Application US/09335325
/ Patent No. US20020164717A1
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ APPLICANT: O'Reilly, Micheal
/ Cao, Yihai
/ Sim, B. Kim Lee
/ TITLE OF INVENTION: Angiostatin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S.
/ ZIP: 30303-1769
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/335,325
/ FILING DATE: 17-Jun-1999
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 20:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 78 amino acids
/ TYPE: amino acid
```

```

; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Porcine
; IMMEDIATE SOURCE:
; CLONE: K3
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-335-325-20
Query Match          57.2%; Score 299; DB 10; Length 78;
Best Local Similarity 56.4%; Pred. No. 6.9e-26;
Matches 44; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGTFSTVTGTCQSWSMTPHRRHQRTPENYPNDGLTMNYCRNPDA DTGFW 63
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CLKGRGENTRGTVSVTASGHTCQWQAQSPHKHNETPENFCCKNLEENYCRNP DGTAPW 60

QY 64 CFTTDPISIRWEYCNLTRC 81
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CYITDSEVRWDYCKIPSC 78

RESULT 14
US-10-131-241-20
; Sequence 20, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Porcine
US-10-131-241-20
Query Match          57.2%; Score 299; DB 15; Length 78;
Best Local Similarity 56.4%; Pred. No. 6.9e-26;
Matches 44; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGTFSTVTGTCQSWSMTPHRRHQRTPENYPNDGLTMNYCRNPDA DTGFW 63
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CLKGRGENTRGTVSVTASGHTCQWQAQSPHKHNETPENFCCKNLEENYCRNP DGTAPW 60

QY 64 CFTTDPISIRWEYCNLTRC 81
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CYITDSEVRWDYCKIPSC 78

RESULT 15
US-09-335-325-18
; Sequence 18, Application US/09335325
; Patent No. US20020164717A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; O'Reilly, Micheal
; Cao, Yihai

```

```

; Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: K3
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-335-325-18
Query Match          54.5%; Score 285; DB 10; Length 78;
Best Local Similarity 55.1%; Pred. No. 2.5e-24;
Matches 43; Conservative 10; Mismatches 25; Indels 0; Gaps 0;

QY 4 CYHNGQSYRGTFSTVTGTCQSWSMTPHRRHQRTPENYPNDGLTMNYCRNPDA DTGFW 63
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CLKGTGENTRGVAVTVSGHTCQWQAQTPHNTNPENFCKNLDENYCRNP DCKRAPW 60

QY 64 CFTTDPISIRWEYCNLTRC 81
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CHTNSQVRWEYCKIPSC 78

Search completed: September 29, 2003, 08:45:29
Job time : 62 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
Run on: September 29, 2003, 08:30:58 ; Search time 39 seconds
(without alignments)
219.462 Million cell updates/sec

Title: US-10-088-548-6
Perfect score: 523
Sequence: 1 VRQYHGNGSYRGTFSTTV.....SIRWEYCNLTRGSDTEGVV 89

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 30955

Minimum DB seq length: 0
Maximum DB seq length: 89

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR/J6:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	354	67.7	89	2 A60140	plasmin (EC 3.4.21.7)
2	64	12.2	64	2 JX0060	short neurotoxin 2
3	61.5	11.8	61	1 N1R12	short neurotoxin 1
4	54	10.3	67	2 S24216	Ig kappa chain - m
5	52	9.9	76	2 T34687	hypothetical prote
6	52	9.9	81	2 A41441	intestinal trefoil
7	50	9.6	87	2 B60549	kappa-3 neurotoxin
8	49.5	9.5	77	2 T14395	proteinase inhibit
9	49.5	9.5	88	2 S41030	hypothetical prote
10	49	9.4	55	2 S58892	H+-transporting tw
11	49	9.4	82	2 S23703	capsid gene 5'-reg
12	48	9.2	72	2 S41611	high-potential iro
13	48	9.2	81	2 AC0142	conserved hypothet
14	47.5	9.1	84	2 A26667	p52 protein precu
15	47.5	9.1	85	2 G69176	hypothetical prote
16	47	9.0	60	1 N1EX1	short neurotoxin 1
17	46.5	8.9	79	2 T11380	NADH2 dehydrogenas
18	46	8.8	73	2 AF0060	hypothetical prote
19	46	8.8	77	2 D86401	protein T22C5.7 [i
20	46	8.8	77	2 T14930	hypothetical prote
21	46	8.8	80	2 S23963	intestinal trefoil
22	45.5	8.7	77	2 S20578	protease inhibitor
23	45	8.6	52	2 S63324	hypothetical prote
24	45	8.6	59	1 T6EP5J	mambin - eastern J
25	45	8.6	60	1 N1AT1	short neurotoxin 1
26	45	8.6	60	1 N1AT1F	short neurotoxin 1
27	45	8.6	60	2 A27580	short neurotoxin 1
28	45	8.6	64	2 E32747	hypothetical prote
29	45	8.6	77	2 S35815	envelope protein -

30	45	8.6	87	1 N1KF2U	long neurotoxin 2
31	44.5	8.5	60	2 A54369	fibrinogen recepto
32	44.5	8.5	62	2 B25866	short neurotoxin 1
33	44.5	8.5	71	2 T17765	hypothetical prote
34	44.5	8.5	86	2 T34975	hypothetical prote
35	44	8.4	61	1 N1R11	short neurotoxin 1
36	44	8.4	67	2 D81915	hypothetical prote
37	44	8.4	88	2 G69136	ribosomal protein
38	43.5	8.3	54	2 T35239	very hypothetical
39	43.5	8.3	62	2 F25866	short neurotoxin c
40	43.5	8.3	74	2 T31295	hypothetical prote
41	43.5	8.3	77	2 B64433	proteinase inhibit
42	43	8.2	35	2 JN0361	insect toxin 15 -
43	43	8.2	35	2 S06667	toxin-like protein
44	43	8.2	60	1 N1AY1	short neurotoxin A
45	43	8.2	68	2 S26474	Ig kappa chain V r

ALIGNMENTS

RESULT 1
A60140
plasmin (EC 3.4.21.7) precursor - chicken (fragment)
M:Alternate names: plasminogen
C:Species: Gallus gallus (Chicken)
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 16-Jul-1999
C:Accession: A60140
R:Gyenes, M.; Patthy, L.
Biochim. Biophys. Acta 832, 326-330, 1985
A:Title: The kringle 4 domain of chicken plasminogen.
R:Reference number: A60140; MUID:86077796; PMID:4074753
A:Accession: A60140
A:Molecule type: protein
A:Residues: 1-89 <GYE>
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homo.
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinas.
P:6-83/Domain: kringle homology <RG>
P:6-83/27-66,55-78/Disulfide bonds: #status predicted
P:39/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 67.7%; Score 354; DB 2; Length 89;
Best Local Similarity 67.9%; Pred. No. 2.1e-29;
Matches 55; Conservative 11; Mismatches 15; Indels 0; Gaps 0;
QY 3 OCYHGNGSYRGTFSTTVGRTCSQSSMTPHRQTPENYNDGLTMYCRNPDATGP 62
DB 5 ECIQGVSVRGASTFTTGKCKQAWNSPHRNKTESHFPNADLRQNYCRNPDADSP 64
QY 63 WCFTTDFPSIRWEYCNLTRGSD 83
DB 65 WCYTTDPSVRWEYCNLKRCS 85

RESULT 2
JX0060
short neurotoxin - banded krait
C:Species: Bungarus fasciatus (banded krait)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 18-Jun-1993
R:Liu, C.S.; Chen, J.P.; Chang, C.S.; Lo, T.B.
J. Biochem. 105, 93-97, 1989
A:Title: Amino acid sequence of a short chain neurotoxin from the venom of banded krai
A:Reference number: JX0060; MUID:89291813; PMID:2738048
C:Accession: JX0060
A:Molecule type: protein
A:Residues: 1-64 <LIU>
C:Superfamily: snake toxin
C:Keywords: postsynaptic neurotoxin

Query Match 12.2%; Score 64; DB 2; Length 64;
Best Local Similarity 26.7%; Pred. No. 7.2;
Matches 16; Conservative 7; Mismatches 27; Indels 10; Gaps 1;

```
QY 34 HRHQTPEPNYPNDG-----LTMNYCRNPADTGPWCFTTDPDSIRWEYCNLTRCS 83
Db 5 NQOQSTPEQPTNGQCYIKTCQKNWTHRGSRTRDRCGCPKYKPGINLRCKCKDKONE 64

RESULT 3
NLR12
short neurotoxin 2 - ringhals
N/Alternate names: neurotoxin IV
C/Species: Hemachatus haemachatus, Sepedon haemachatus (ringhals)
C/Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 23-Aug-1996
C/Accession: A01701
R/Stridom, A.J.C.; Botes, D.P.
J. Biol. Chem. 246, 1341-1349, 1971
A/Title: Snake venom toxins. Purification, properties, and complete amino acid sequence
A/Reference number: A92073; MUID:71116407; PMID:5545078
A/Accession: A01701
A/Molecule type: protein
A/Residues: 1-61 <SWR>
C/Superfamily: snake toxin
C/Keywords: neurotoxin; venom
F:3-23,17-40,42-53,54-59/Disulfide bonds: #status predicted

Query Match 11.8%; Score 61.5; DB 1; Length 61;
Best Local Similarity 26.1%; Pred. No. 12;
Matches 23; Conservative 8; Mismatches 20; Indels 37; Gaps 5;

QY 3 QCYHGNGQSVRGFTFTVTGRTG-----QSWSSMTHRHQRTPEPNYPNDGLTMNYCR 54
Db 2 ECH--NQOS-----SQTPTQCPGETNCKYKQWSD--HRGSRT----- 36

QY 55 NPADATGPWCFTTDPDSIRWEYCNLTRCS 82
Db 37 -----ERCGCPYKPGIKLKCCYTDRCN 60

RESULT 4
S24216
Iq kappa chain - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C/Accession: S24216
R/Shimizu, T.; Iwasato, T.; Yamagishi, H.
J. Exp. Med. 173, 1065-1072, 1991
A/Title: Deletions of immunoglobulin C(kappa) region characterized by the circular excis
A/Reference number: S24214; MUID:91217618; PMID:1902500
A/Accession: S24216
A/Status: preliminary; translation not shown
A/Molecule type: DNA
A/Residues: 1-67 <SHI>
A/Cross-references: EMBL:X58221; NID:953719; PIDN:CAA41185.1; PID:9930196
C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: heterotetramer; immunoglobulin

Query Match 10.3%; Score 54; DB 2; Length 67;
Best Local Similarity 39.5%; Pred. No. 80;
Matches 15; Conservative 1; Mismatches 18; Indels 4; Gaps 1;

QY 1 VRQCYHGNGQSVRGFTFTVTGRTG-----TTVTGRTGQSWSSMTPH 34
Db 20 VRFGSGSGGTYSLSITSRMEADATYCCQWSYPPH 57

RESULT 5
T34687
hypothetical protein SC1B5.09c SC1B5.09c - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C/Accession: T34687
R/Harris, D.; Taylor, K.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1998
A/Reference number: 221553
```

```
A/Accession: T34687
A/Status: preliminary; translated from GS/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-76 <HAR>
A/Cross-references: EMBL:AL023517; PIDN:CAA18983.1; GSPDB:GM00070; SCOEDB:SC1B5.09c
A/Experimental source: strain A3(2)
C/Genetics:
A/Gene: SCOEDB:SC1B5.09c
```

```
Query Match 9.9%; Score 52; DB 2; Length 76;
Best Local Similarity 27.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 2; Mismatches 24; Indels 28; Gaps 4;

QY 34 HRHQTPEPNYPNDGLTM---NYC-----RNPADTGPWC-----FTTD 68
Db 5 HGDREPE---PDGAGEPGYGLIDAPPPKADGPPFAECVQCRPEPTPEPSYKGLIC 61

QY 69 PSIRWEYCNLTRCS 82
Db 62 PVCEWQEAQRTACS 75

RESULT 6
A41441
intestinal trefoil factor precursor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 16-Jul-1999
C/Accession: A41441
R/Suenori, S.; Lynch-Devaney, K.; Podolsky, D.K.
Proc. Natl. Acad. Sci. U.S.A. 88, 11017-11021, 1991
A/Title: Identification and characterization of rat intestinal trefoil factor: tissue
A/Reference number: A41441; MUID:92107881; PMID:1763017
A/Accession: A41441
A/Molecule type: mRNA
A/Residues: 1-81 <SUE>
A/Cross-references: GB:M60826; NID:9207446; PIDN:AAA42270.1; PID:9207447
A/Note: the authors translated the codon ACA for residue 8 as Ile
C/Superfamily: secretory protein xP1; trefoil homology
C/Keywords: intestine
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-22/Domain: propeptide #status predicted <PRO>
F:23-81/Product: intestinal trefoil factor #status predicted <IAT>
F:33-73/Domain: trefoil homology <TRF>
F:33-59,43-58,53-70/Disulfide bonds: #status predicted
```

```
Query Match 9.9%; Score 52; DB 2; Length 81;
Best Local Similarity 28.2%; Pred. No. 1.5e+02;
Matches 20; Conservative 8; Mismatches 17; Indels 26; Gaps 6;

QY 16 FSTT-----VTGRTG--QSWSSMTPHR-----HQRTPEPNYPNDGLTMNYCRN----- 55
Db 6 FWTLLLVLAGSSCKAQEFVGLSPSQCAPTNVRVDCNYPT--VTSEQNNRGCCFDSS 63

QY 56 -PDADTGPWC 65
Db 64 IPNV---PWC 71
```

```
RESULT 7
B60549
kappa-3 neurotoxin precursor - many-banded krait
N/Alternate names: kappa-3 bungarotoxin; kappa-neurotoxin CRI
C/Species: Bungarus multicinctus (many-banded krait)
C/Date: 17-Apr-1993 #sequence_revision 18-Jun-1993 #text_change 17-Mar-2000
C/Accession: S08398; B60549
R/Dansey, J.M.; Garnier, J.M.
Nucleic Acids Res. 18, 1050, 1990
A/Title: cDNA deduced amino-acid sequences of two novel kappa-neurotoxins from Bungar
A/Reference number: S08398; MUID:90192091; PMID:2315018
A/Accession: S08398
A/Molecule type: mRNA
A/Residues: 1-87 <DAN>
A/Cross-references: EMBL:X51413; NID:962481; PIDN:CAA35775.1; PID:g62482
```

Query Match 9.5%; Score 49.5; DB 2; Length 88;
Best Local Similarity 29.4%; Pred. No. 3e+02;
Matches 15; Conservative 7; Mismatches 14; Indels 15; Gaps 2;

high-potential iron sulfur protein isoform 1 - Ectothiorhodospira

N;Alternate names: HiIP protein

RESULT 15

G69176
Hypothetical protein MTH58 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: G69176
R:Smith, D.R.; Doucet-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Oiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.;
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: G69176
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-85 <MTH>
A:Cross-references: GB:AE000798; GB:AE000656; NID:g2621094; PIDN:AAB84565.1; PID:g262109
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH58

```

Query Match          9.18; Score 47.5; DB 2; Length 85;
Best Local Similarity 25.0%; Pred. No. 4.7e-02;
Matches 19; Conservative 12; Mismatches 26; Indels 19; Gaps 4;

QY      18 TTVTGRTCOSGSSMPHHQRTPEX---YPNDGLNNYCRNPD-ADTGPCVTTDPSIRW 73
      |||:  :  |||:  :  |||:  :  |||:  :  |||:  :  |||:  :
Db       8 TGITGK-----QRPIDGVVPEDSYLVTWPKGWLFGP-VTARTEIKY 52

QY      74 EYCNLTRECSDEGTGVV 89
      :  | :  :  |||:
Db      53 GRPHQTKLTLETSDTVI 68

```

Search completed: September 29, 2003, 08:36:31
Job time : 41 secs

Result No.	Query %			ID	Description
	Score	Match	Length		
1	64	12.2	64	1	NXS1_BUNFA
2	61.5	11.8	61	1	NXS2_HEMHA
3	50.5	9.7	81	1	POC3_SVRVU
4	50.5	9.7	84	1	POC3_OLEEU
5	50	9.6	87	1	NXL3_BUNMU
6	49	9.4	55	1	ATP8_LATCH
7	48	9.2	72	1	HP11_ECTVA
8	47.5	9.1	84	1	TFE1_HUMAN
9	47	9.0	60	1	NXS1_ENHSC
10	46	8.8	81	1	TFE3_RAT
11	45.5	8.7	37	1	SC1P_MESTA
12	45.5	8.7	77	1	THG1_ARAYE
13	45	8.6	52	1	YN78_YEAST
14	45	8.6	59	1	MAMB_DENJA
15	45	8.6	60	1	NXS1_ASTST
16	45	8.6	60	1	TXW3_NAJHA
17	45	8.6	61	1	TXWA_NAJHA
18	45	8.6	67	1	OTX2_RAT
19	45	8.6	70	1	NXL1_BUNMU
20	44.5	8.5	60	1	DISC_ORNMO
21	44.5	8.5	62	1	NXS2_LATCO
22	44.5	8.5	62	1	NXS4_LATCR
23	44	8.4	61	1	NXS1_HEMHA
24	44	8.4	88	1	RL34_METH
25	43.5	8.3	62	1	NXSC_LATCR
26	43.5	8.3	77	1	THG2_ARATH
27	43.5	8.3	82	1	CXMB_CNMR
28	43.5	8.3	86	1	TXW2_NAJAT
29	43	8.2	35	1	SCX1_BUTSI
30	43	8.2	60	1	NXS1_AIPLA
31	43	8.2	60	1	NXS3_AIPLA
32	43	8.2	60	1	SCX3_MESTA
33	43	8.2	65	1	YSF1_RUILO


```

Query Match          9.4%; Score 49; DB 1; Length 55;
Best Local Similarity 34.1%; Pred. No. 56;
Matches 14; Conservative 5; Mismatches 16; Indels 6; Gaps 3;

QY      27 SW--SSMPHGHQRTPEPTNPDGLTMNYCR-NPDADIGPW 63
      || :||| :|| :|| :|| :|| :|| :|| :||
Db       16 SWLFLTMLPSKTQL--HTFPNMPSTQNMCKQEPFPTWFW 54

RESULT 7
HP11_ECTVA
ID      HP11_ECTVA      STANDARD;      PRT;      72 AA..
AC      P38941;
DT      01-FEB-1995 (Rel. 31, Created)
DT      01-FEB-1995 (Rel. 31, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      High potential iron-sulfur protein isozyme 1 (HiP1 1).
GN      HIP1.
OS      Ectothiorhodospira vacuolata.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales;
OC      Ectothiorhodospiraceae; Ectothiorhodospira.
OX      NCBI TaxID=1054;

```


RX MEDLINE=6832627; PubMed=3261981;
 RA Mori K., Fujii R., Kida N., Ohka M., Hayashi K.;
 RT "Identification of a polypeptide secreted by human breast cancer
 RL cells (MCF-7) as the human estrogen-responsive gene (ps2) product.";
 RN Biochem. Biophys. Res. Commun. 155:366-372(1988).
 RP
 RX SEQUENCE OF 25-84.
 RA MEDLINE=89119134; PubMed=3146413;
 RT Rio M.-C., Lepage P., Diemunsch P., Roitsch C., Chambon P.;
 RL "Primary structure of human protein ps2.";
 RN C. R. Acad. Sci., III, Sci. Vie 307:825-831(1988).
 RP
 RX BIOSYNTHESIS, AND SECRETION.
 RA MEDLINE=88290699; PubMed=3041593;
 RT Rio M.-C., Bellocq J.-P., Daniel J.-Y., Tomasetto C., Lathe R.,
 RL Chenard M.P., Batzenslager A., Chambon P.;
 RN "Breast cancer-associated ps2 protein: synthesis and secretion by
 RT normal stomach mucosa.";
 RL Science 241:705-708(1988).
 RP
 RX STRUCTURE BY NMR.
 RA MEDLINE=96085149; PubMed=8521850;
 RT Polshakov V.I., Frenkiel T.A., Westley B.R., Chadwick M.P.,
 RL May F.E.B., Carr M.D., Feeney J.;
 RN "NMR-based structural studies of the pNR-2/ps2 single domain trefoil
 RT peptide. Similarities to porcine spasmodic peptide and evidence for
 RL a monomeric structure.";
 RN Eur. J. Biochem. 233:847-855(1995).
 RP
 RX STRUCTURE BY NMR.
 RA MEDLINE=97250379; PubMed=9096235;
 RT Polshakov V.I., Williams M.A., Gargaro A.R., Frenkiel T.A.,
 RL Westley B.R., Chadwick M.P., May F.E.B., Feeney J.;
 RN "High-resolution solution structure of human pNR-2/ps2: a single
 RT trefoil motif protein.";
 RL J. Mol. Biol. 267:418-432(1997).
 RP
 RX VARIANTS CANCER ILE-32; ASP-34; LYS-37; ILE-46 AND VAL-55.
 RA MEDLINE=20440596; PubMed=10982763;
 RT Park W.-S., Oh R.-R., Park J.-Y., Lee J.-H., Shin M.-S., Kim H.-S.,
 RL Lee H.-K., Kim Y.-S., Kim S.-Y., Lee S.-H., Yoo N.-J., Lee J.-Y.;
 RN "Somatic mutations of the trefoil factor family 1 gene in gastric
 RT cancer.";
 RL Gastroenterology 119:691-698(2000).
 CC
 CC -1- FUNCTION: Stabilizer of the mucous gel overlying the
 CC gastrointestinal mucosa that provides a physical barrier against
 CC various noxious agents.
 CC
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC
 CC -1- TISSUE SPECIFICITY: FOUND IN STOMACH; IS ALSO EXPRESSED IN BREAST
 CC CANCER BUT NOT IN NORMAL TISSUE. IT IS REGULATED BY ESTROGEN IN
 CC MCF-7 CELLS. STRONG EXPRESSION FOUND IN NORMAL GASTRIC MUCOSA AND
 CC IN THE REGENERATIVE TISSUES SURROUNDING ULCEROUS LESIONS OF
 CC GASTROINTESTINAL TRACT.
 CC
 CC -1- DISEASE: EXPRESSED IN A NUMBER OF CARCINOMAS INCLUDING BREAST
 CC (508), PANCREAS (70-75%), STOMACH (50-55%), AND LARGE BOWEL (58%).
 CC
 CC -1- SIMILARITY: Contains 1 P-type (trefoil) domain.
 CC
 CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
 CC WWW="http://www.infobiogen.fr/services/Chromancer/Genes/TF1ID201.html".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X00474; CAA25155.1; -;
 CC EMBL; X05030; CAA28695.1; -;
 CC EMBL; X05321; CAA28695.1; JOINED.
 CC EMBL; X05322; CAA28695.1; JOINED.
 CC EMBL; X52003; CAA36254.1; -;
 CC EMBL; M12075; AAM52402.1; -;

DR EMBL; AB038162; BAB13729.1; -;
 DR EMBL; AP001746; BAA95532.1; -;
 DR EMBL; BC032811; AAH32811.1; -;
 DR PIR; A26667; A26667.
 DR PDB; 1PS2; 07-JUL-97.
 DR PDB; 1HI7; 09-APR-01.
 DR Genew; HGNC:11755; TFF1.
 DR MIM; 113710; -;
 DR GO; GO:0005975; P:carbohydrate metabolism; TAS.
 DR GO; GO:0008151; P:cell growth and/or maintenance; TAS.
 DR InterPro; IPR000519; P:trefoil.
 DR Pfam; PF00088; trefoil.1.
 DR PRINTS; PR00680; PTFREFOIL.
 DR SMART; SM00018; PD; 1.
 DR PROSITE; PS00025; P_TREFOIL; 1.
 DR Growth factor; Signal; Disease mutation; 3D-structure.
 FT SIGNAL 1 24
 FT CHAIN 25 84 TREFOIL FACTOR 1.
 FT DOMAIN 30 71 P-TYPE.
 FT DISULFID 31 57 BY SIMILARITY.
 FT DISULFID 41 56 BY SIMILARITY.
 FT DISULFID 51 68 BY SIMILARITY.
 Query Match 9.1%; Score 47.5; DB 1; Length 84;
 Best Local Similarity 31.0%; Pred. No. 1.3e+02;
 Matches 18; Conservative 6; Mismatches 19; Indels 15; Gaps 6;
 Qy 14 GTFTTIVGRTCSWSSMTPHRTPEYVNDGLTMYCRN-----PDADTG-PWCF 65
 Db 21 GTLAERQT-ETC-----TVAPERQNC--GFP--GVTFQCANKGCCFDDTVRGVPWCF 69
 RESULT 9
 NXSL_ENHSC STANDARD; PRT; 60 AA.
 ID NXSL_ENHSC
 AC P25492; P01436;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Short neurotoxin 1 (Toxin 4).
 OS Enhydrina schistosoma (Beaked sea snake) (Common sea snake), and
 OS Lepadotis hardwickii (Hardwick's sea snake).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidodactylidae; Squamata; Scleroglossa; Serpentes; Colubroides;
 OC Elapidae; Hydrophiinae; Enhydrina.
 OX NCBI_TaxID=8682, 8781;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=E. schistosoma; TISSUE=Venom;
 RX MEDLINE=73127342; PubMed=4676045;
 RA Fryklund L., Eaker D., Karlsson E.;
 RT "Amino acid sequences of the two principal neurotoxins of Enhydrina
 RT schistosoma venom.";
 RL Biochemistry 11:4633-4640(1972).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=L. hardwickii; TISSUE=Venom;
 RX MEDLINE=77246775; PubMed=891966;
 RA Fox J.W., Elzinga M., Tu A.T.;
 RT "Amino acid sequence of a snake neurotoxin from the venom of Lepadotis
 RT hardwickii and the detection of a sulfhydryl group by laser Raman
 RT spectroscopy.";
 RL FEBS Lett. 80:217-220(1977).
 CC -1- FUNCTION: PRODUCES PERIPHERAL PARALYSIS BY BLOCKING NEUROMUSCULAR
 CC TRANSMISSION AT THE POSTSYNAPTIC SITE. BINDS TO THE NICOTINIC
 CC ACETYLCHOLINE RECEPTOR.
 CC
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC
 CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC
 CC -1- MISCELLANEOUS: LD(50) IS 0.2 MG/KG BY INTRAVENOUS INJECTION FOR
 CC E. SCHISTOSOMA.
 CC
 CC -1- SIMILARITY: Belongs to the snake toxin family.
 DR PIR; A01705; NIEY1.
 DR HSSP; P01426; INEA.

RESULT 11

SCIT_MESTA STANDARD; PRT; 37 AA.
 ID P81761;
 AC 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lepidopteran-selective toxin (Butait).
 OS Mesobuthus tamulus (Eastern Indian scorpion) (Buthus tamulus).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
 OC Buthoidea; Mesobuthus.
 OX NCBI_TaxID=34647;
 RN [1]
 RP SEQUENCE, FUNCTION, SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND MASS
 RP SPECTROMETRY.
 RC TISSUE-Venom.
 RX PubMed=11762289;
 RA Wudayagiri R., Inceoglu B., Herrmann R., Derbel M., Choudary P.V.,
 RA Hammock B.D.;
 RT "Isolation and characterization of a novel lepidopteran-selective
 RT toxin from the venom of South Indian red scorpion, Mesobuthus
 RT tamulus.";
 RL BMC Biochem. 2:16-16(2001).
 CC -1- FUNCTION: Induces flaccid paralysis in H.virens larvae. Is not
 CC toxic to S.falculata larvae or mice.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC -1- MASS SPECTROMETRY: MW=3856.7; METHOD-Electrospray.
 CC -1- SIMILARITY: BELONGS TO THE SHORT SCORPION TOXIN FAMILY.
 DR GO: GO:0005576; C:extracellular; IDA.
 DR GO: GO:0019870; F:potassium channel inhibitor activity; NAS.
 DR GO: GO:0009405; P:pathogenesis; IDA.
 DR Pfam: PF05294; toxin_5; 1.
 KW Toxin; Neurotoxin.
 FT DISULFID 2 19 BY SIMILARITY.
 FT DISULFID 5 27 BY SIMILARITY.
 FT DISULFID 16 32 BY SIMILARITY.
 FT DISULFID 20 34 BY SIMILARITY.
 SQ SEQUENCE 37 AA; 3864 MW; FD36529B5FA7CCD9 CRC64;

Query Match 8.7%; Score 45.5; DB 1; Length 37;
 Best Local Similarity 42.3%; Pred. No. 87;
 Matches 11; Conservative 3; Mismatches 5; Indels 7; Gaps 2;

QY 61 GPWCFTDPDSRWYCNLCRCSDTEG 86
 II IIIII : : : : :
 Db 3 GP-CFTDPQTQ-----AKCSECG 21

RESULT 12

THG1_ARATH STANDARD; PRT; 77 AA.
 ID THG1_ARATH
 AC Q39182; Q42011;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gamma-thionin homolog At2g02100 precursor.
 GN AT2G02100 OR F504.13.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. C24; TISSUE=Flower buds;
 RA Yu D.Y., Quigley F., Mache R.;
 RT "Isolation and expression of a cDNA encoding protease inhibitor
 RT precursor".
 RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;

RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
 RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
 RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
 RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
 RA Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 402:761-768(1999).
 RN [3]
 RP SEQUENCE OF 1-75 FROM N.A.
 RC STRAIN=cv. C24; TISSUE=Flower buds;
 RA Mache R., Quigley F., Thomas F., Yu D.Y.;
 RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: BELONGS TO THE PLANT DEPENSIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X69139; CAA48892.1; -;
 DR EMBL; AC005936; AAC97223.1; -;
 DR EMBL; Z18455; CAA79189.1; -;
 DR PIR; S30578; S30578.
 DR HSP; P20230; LGFT
 DR InterPro: IPR002118; Gamma-thionin.
 DR InterPro: IPR003614; Knot1.
 DR Pfam: PF00304; Gamma-thionin; 1.
 DR ProDom: PD002594; G_Purothionin; 1.
 DR SMART: SM00505; Knot1; 1.
 DR PROSITE: PS00940; GAMMA_THIONIN; 1.
 DR Plant defense; Signal; Multigene family.
 KW SIGNAL
 FT CHAIN 1 30 POTENTIAL.
 FT DISULFID 31 77 GAMMA-THIONIN HOMOLOG AT2G02100.
 FT DISULFID 33 77 BY SIMILARITY.
 FT DISULFID 44 64 BY SIMILARITY.
 FT DISULFID 50 71 BY SIMILARITY.
 FT DISULFID 54 73 BY SIMILARITY.
 SQ SEQUENCE 77 AA; 8524 MW; EC04E660C3610965 CRC64;
 Query Match 8.7%; Score 45.5; DB 1; Length 77;
 Best Local Similarity 35.7%; Pred. No. 1,9e+02;
 Matches 15; Conservative 3; Mismatches 17; Indels 7; Gaps 2;
 QY 14 GFSTTVTGRTCSWSMTPHRHQRTPEPNPDGLTMNVCN 55
 | | | | | : : : : :
 Db 22 GMGPVTVEARTCESQS-----HFKGTGCVSASN--CANVCHN 56
 RESULT 13
 YN78_YEAST
 ID YN78_YEAST STANDARD; PRT; 52 AA.
 AC P53820;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 6.0 kDa protein in COS1 5' region.
 GN YNL338W OR N0170.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Obermaier B., Piravandi E., Rinke M.;

```

RL Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC -I- SIMILARITY: TO YEAST YHR217C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z71614; CAA96274.1; -.
DR EMBL; Z71613; CAA96273.1; -.
DR PIR; S63324; S63324.
DR SGD; S0005282; YNL336W.
KW Hypothetical protein.
SQ SEQUENCE 52 AA; 5951 MW; C1E4066D43E057A1 CRC64;

Query Match      8.6%; Score 45; DB 1; Length 52;
Best Local Similarity 28.9%; Pred. No. 1.4e+02;
Matches 13; Conservative 5; Mismatches 17; Indels 10; Gaps 1;

QY 23 RTQCSWSSM-----TPRHQRTPEPNYNDGLTMYNCRNPD 57
   ||: ||| || | || | || | || | || |
Db 4 RPLCTPSMQYSDIYPTPTTTHRTHTTPHPHPHTHTHNP 48

RESULT 14
NAME_DENJA STANDARD; PRT; 59 AA.
AC P28375;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mambin (Glycoprotein IIB-IIIa antagonist) (Platelet aggregation
DE inhibitor) (dendroaspin).
OS Dendroaspis jamestoni kamosae (Eastern Jameson's mamba).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Dendroaspis.
OX NCBI_TaxID=8619;
RN [1]
RP TISSUE=Venom;
RC MEDLINE=92273586; PubMed=1591238;
RA McDowell R.S., Dennis M.S., Louie A., Shuster M., Mulkerrin M.G.,
RA Lazarus R.A.;
RT "Mambin, a potent glycoprotein IIB-IIIa antagonist and platelet
RT aggregation inhibitor structurally related to the short
RT neurotoxins.";
RL Biochemistry 31:4766-4772(1992).
RN [2]
RP STRUCTURE BY NMR.
RX MEDLINE=95360732; PubMed=7634091;
RA Sutcliffe M.J., Jaseja M., Hyde E.I., Lu X., Williams J.A.;
RT "Three-dimensional structure of the RGD-containing neurotoxin
RT homologue dendroaspin.";
RL Nat. Struct. Biol. 1:802-812(1994).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE=95112851; PubMed=7813476;
RA Jaseja M., Lu X., Williams J.A., Sutcliffe M.J., Kakkar V.V.,
RA Parslow R.A., Hyde E.I.;
RT "1H-NMR assignments and secondary structure of dendroaspin, an RGD-
RT containing glycoprotein IIB-IIIa (alpha IIB-beta 3) antagonist with a
RT neurotoxin fold.";
RL Eur. J. Biochem. 226:861-869(1994).
CC -I- FUNCTION: INHIBITS ADP-INDUCED PLATELET AGGREGATION AND INHIBITS
CC THE BINDING OF PURIFIED PLATELET FIBRINOGEN RECEPTOR GPIIb-IIIa
CC TO IMMOBILIZED FIBRINOGEN.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -I- SIMILARITY: Belongs to the snake toxin family.

```

```

DR PIR; A42561; T6EP5J.
DR PDB; 1DRS; 20-DEC-94.
DR InterPro; IPR003571; Snake_toxin.
DR Pfam; PF00087; toxin; 1.
DR ProDom; PD000206; Snake_toxin; 1.
DR PROSITE; PS00272; SNAKE_TOXIN; FALSE_NEG.
KW Blood coagulation; Cell adhesion; Toxin; 3D-structure.
FT DISULFID 3 22
FT DISULFID 17 37
FT DISULFID 39 51
FT DISULFID 52 57
FT SITE 43 45 CELL ATTACHMENT SITE (POTENTIAL).
FT STRAND 2 4
FT TURN 8 9
FT STRAND 14 16
FT STRAND 22 24
FT TURN 28 31
FT STRAND 36 37
FT STRAND 50 52
SQ SEQUENCE 59 AA; 6754 MW; 20DDC6A5D9DF9E41 CRC64;

Query Match      8.6%; Score 45; DB 1; Length 59;
Best Local Similarity 28.2%; Pred. No. 1.6e+02;
Matches 22; Conservative 8; Mismatches 28; Indels 20; Gaps 5;

QY 2 RQYHGNGQSYRGTFSTTVTGRTCQSWSSMTPHRHQRTPEPNYNDGLTMYNCRNPDAD-T 60
   ||: ||| || | || | || | || | || | || | || |
Db 1 RICYNHLG-----TKPPTETCQEDSC---YKNWTFDNIIRGCG---CFTPRGDMP 47

QY 61 GPWCFTTDFSIWYCNL 78
   ||: || || | || |
Db 48 GPYCCESDR-----CNL 59

RESULT 15
NAME_LASTST STANDARD; PRT; 60 AA.
AC P01438; P10461; P19005;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Short neurotoxin I (Toxin A).
OS Astrotia stokesi (Stokes's sea snake) (Disteira stokesi),
OS Hydrophis ornatus (Reef sea snake), and
OS Acalyptophis peronii (Sea snake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Hydrophiinae; Astrotia.
OX NCBI_TaxID=8680, 8685, 8676;
RN [1]
RP SEQUENCE.
RC SPECIES=A.stokesi; TISSUE=Venom;
RX MEDLINE=79123997; PubMed=743209;
RA Maeda N., Tamiva N.;
RT "Three neurotoxins from the venom of a sea snake Astrotia stokesii,
RT including two long-chain neurotoxic proteins with amidated
RT C-termini.";
RL Biochem. J. 175:507-517(1978).
RN [2]
RP COMPOSITION OF TRYPTIC PEPTIDES.
RC SPECIES=H.ornatus; TISSUE=Venom;
RX MEDLINE=83308533; PubMed=6615431;
RA Tamiya N., Maeda N., Cogger H.G.;
RT "Neurotoxins from the venoms of the sea snakes Hydrophis ornatus and
RT Hydrophis lapemoides.";
RL Biochem. J. 213:31-38(1983).
RN [3]
RP SEQUENCE (MAJOR VARIANT).
RC SPECIES=A.peronii; TISSUE=Venom;
RX MEDLINE=88132805; PubMed=3341735;
RA Mori N., Tu A.T.;
RT "Isolation and primary structure of the major toxin from sea snake,
RT Acalyptophis peronii, venom.";

```

```

RL Arch. Biochem. Biophys. 260:10-17(1988).
RN [4]
RP SEQUENCE (MINOR VARIANT).
RC SPECIES=A. peronii; TISSUE=Venom;
RX MEDLINE=89076487; PubMed=3202959;
RA Mori N., Tu A.T.;
RT "amino-acid sequence of the minor neurotoxin from Acalyptophis
RT peronii venom.";
RL Biol. Chem. Hoppe-Seyler 369:521-526(1988).
CC -!- FUNCTION: PRODUCES PERIPHERAL PARALYSIS BY BLOCKING NEUROMUSCULAR
CC TRANSMISSION AT THE POSTSYNAPTIC SITE. BINDS TO THE NICOTINIC
CC ACETYLCHOLINE RECEPTOR.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- MISCELLANEOUS: LD(50) IS 0.170 MG/KG BY INTRAVENOUS INJECTION FOR
CC A. PERONII MINOR TOXIN, 0.125 MG/KG FOR THE MAJOR TOXIN.
CC -!- MISCELLANEOUS: LD(50) IS 0.13 MG/KG BY INTRAMUSCULAR INJECTION FOR
CC A. STOKESI.
CC -!- SIMILARITY: Belongs to the snake toxin family.
DR PIR; A01706; N1AT1.
DR PIR; A27580; A27580.
DR PIR; B90321; N1AT1F.
DR HSSP; P01426; INEA.
DR InterPro; IPR003571; Snake_toxin.
DR Pfam; PF00087; toxin; 1.
DR ProDom; PD000206; Snake_toxin; 1.
DR PROSITE; PS00272; SNAKE_TOXIN; 1.
KW Toxin; Neurotoxin; Postsynaptic neurotoxin; Multigene family.
FT DISULFID 3 22 BY SIMILARITY.
FT DISULFID 17 39 BY SIMILARITY.
FT DISULFID 41 52 BY SIMILARITY.
FT DISULFID 53 58 BY SIMILARITY.
FT VARIANT 43 43 Q->E (IN A. PERONII MINOR VARIANT).
SQ SEQUENCE 60 AA; 6599 MW; DE41989F9056E59A CRC64;

Query Match 8.6%; Score 45; DB 1; Length 60;
Best Local Similarity 21.8%; Pred. NO. 1.7e+02;
Matches 17; Conservative 12; Mismatches 25; Indels 24; Gaps 4;

Qy 8 NCOSYRGTFSTVTGRTG--QSWSSMTFRRHQRTPEYFNDGLTMYCNPADTGPWCF 65
Db 5 NQSSSQPKTTNCAGNSCYKKTWSD-----HRGT-----IIRGGCGPQVKSG---- 47

Qy 66 TTDPSIRWEYCNLTGSD 83
Db 48 -----IKLECCHTNECNN 60

```

Search completed: September 29, 2003, 08:34:03
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:30:34 ; Search time 94 Seconds
(without alignments)
244.326 Million cell updates/sec

Title: US-10-088-548-6
Perfect score: 523
Sequence: 1 VRQCYHNGSGYRGTFSTTV.....SIRWEYCNLTRCSDTEGTVV 89

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 113503

Minimum DB seq length: 0
Maximum DB seq length: 89

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	223	42.6	53	4	Q9UMI2		Q9um12 homo sapien
2	190	36.3	60	4	Q9UKJ7		Q9ukj7 homo sapien
3	141	27.0	25	4	Q9UD88		Q9ud88 homo sapien
4	127.5	24.4	75	6	Q9BGN9		Q9bgn9 bos taurus
5	59.5	11.4	86	15	Q9ORV4		Q9orv4 human immun
6	56	10.7	87	15	Q9OBV2		Q9obv2 human immun
7	55.5	10.6	87	10	Q8H444		Q8h444 oryza sativ
8	53.5	10.2	85	16	Q8FTS2		Q8fts2 corynebacte
9	52.5	10.0	82	15	Q9YL18		Q9yl18 human immun
10	52.5	10.0	87	12	Q8V5I3		Q8v5i3 human respi
11	52.5	10.0	88	15	Q9ID04		Q9id04 human immun
12	52.5	10.0	89	15	Q9ID03		Q9id03 human immun
13	52	9.9	39	6	Q9MZF7		Q9mfz7 bos taurus
14	52	9.9	46	13	Q9FWJ8		Q9fwj8 gallus gall
15	52	9.9	67	15	Q9QJH2		Q9qjh2 human immun
16	52	9.9	76	16	O69837		O69837 streptomyce

Q9lnd7 cactus viru
Q8bba5 human respi
Q8bba4 human respi
Q9gji4 human immun
Q8v5h6 human respi
Q9dxn2 human immun
Q8h3h3 oryza sativ
Q8v5i6 human respi
Q8bba3 human respi
Q8v5i7 human respi
Q72408 human immun
Q9ibh8 human respi
Q8v5i2 human respi
Q8bba8 human respi
Q9i935 human immun
Q9dxx9 human immun
Q8t0x4 mesobuthus
Q9ezg6 streptococc
Q9z3a6 unidentified
Q76655 human immun
Q8mqk9 schistosoma
Q8vad4 white spot
Q39403 brassica ca
O44062 apicomplexa
Q9iud1 human respi
Q9ibc3 human respi
Q8v5h5 human respi
Q8v5i4 human respi
Q90ab8 human immun

ALIGNMENTS

RESULT 1
Q9UMI2
ID Q9UMI2 PRELIMINARY; PRT; 53 AA.
AC Q9UMI2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE PLG protein (fragment).
GN PLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie P.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; K02921; AAA60123.1; -;
DR HSSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PRO0018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; FALSE_NEG.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle.
FT NON_TER 1
SQ SEQUENCE 53 AA; 5893 MW; 2EE36E715CD40A47 CRC64;
Query Match 42.6%; Score 223; DB 4; Length 53;
Best Local Similarity 82.2%; Pred. No. 1.1e-18;
Matches 37; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VRQCYHNGSGYRGTFSTTVGTTCQSWSMTPHRHQTPEYYPN 45


```

RA Sullivan M., Gwinn M., Clark K.A., Busch M.P.;
RT "Use of the Dual Sensitive/Less-Sensitive (Detuned) EIA Strategy for
RT Targeting Genetic Analysis of HIV-1 to Recently Infected U.S. Blood
RT Donors.;"
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF373767; AAK56387.1; -.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 86
SQ SEQUENCE 86 AA; 9895 MW; 1A0C197C8C58A808 CRC64;

Query Match 11.4%; Score 59.5; DB 15; Length 86;
Best Local Similarity 30.0%; Pred. No. 20;
Matches 21; Conservative 10; Mismatches 18; Indels 21; Gaps 4;

QY 41 ENYPNDGLTM-----NYCRNPDADT-----GPWCFTTD--PSIRWEYCN--LT 79
||: : : : ||: : : : ||: : : : ||: : : : ||
Db 10 ENFYDNAKTIIQLNGSVGIYCERNNNTRKGHIHGKVFYTTDIIIGDKAKCNISLT 69
||: : : : ||: : : : ||: : : : ||: : : : ||

QY 80 RCDSTEGTVV 89
: : || : :
Db 70 KWNDTLKQIV 79

RESULT 6
Q90B02 PRELIMINARY; PRT; 87 AA.
AC Q90B02;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PF64-49;
RA Collins K., Quinones-Mateu M., Wu M., Luze H., Johnson J., Hirsch C.,
RA Toossi Z., Arts E.;
RT "A TB-mediated site Specific Increase in HIV-1 Quasispecies Diversity
RT May Lead to the Observed Greater HIV-1 Heterogeneity in Patients Co-
RT Infected with Tuberculosis.;"
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF383521; AAK98372.1; -.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 87
SQ SEQUENCE 87 AA; 9930 MW; 1D246B516BDD1F2C CRC64;

Query Match 10.7%; Score 56; DB 15; Length 87;
Best Local Similarity 32.7%; Pred. No. 51;
Matches 17; Conservative 9; Mismatches 14; Indels 12; Gaps 4;

QY 45 NDGLTMNYCRNPDADT-----GP----WCFTTD--PSIRWEYCNLRCSDTIE 85
||: : : : ||: : : : ||: : : : ||: : : : ||
Db 16 NESVQIN-CTRPNNNRKSVHIGPGQAFYATGDIIGKIRWAYCNYSRAAWNE 66

RESULT 7
Q8H444 PRELIMINARY; PRT; 87 AA.
AC Q8H444;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE P0407H12.22 protein.
GN P0407H12.22.

RA Sullivan M., Gwinn M., Clark K.A., Busch M.P.;
RT "Use of the Dual Sensitive/Less-Sensitive (Detuned) EIA Strategy for
RT Targeting Genetic Analysis of HIV-1 to Recently Infected U.S. Blood
RT Donors.;"
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF373767; AAK56387.1; -.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 86
SQ SEQUENCE 86 AA; 9895 MW; 1A0C197C8C58A808 CRC64;

Query Match 11.4%; Score 59.5; DB 15; Length 86;
Best Local Similarity 30.0%; Pred. No. 20;
Matches 21; Conservative 10; Mismatches 18; Indels 21; Gaps 4;

QY 41 ENYPNDGLTM-----NYCRNPDADT-----GPWCFTTD--PSIRWEYCN--LT 79
||: : : : ||: : : : ||: : : : ||: : : : ||
Db 10 ENFYDNAKTIIQLNGSVGIYCERNNNTRKGHIHGKVFYTTDIIIGDKAKCNISLT 69
||: : : : ||: : : : ||: : : : ||: : : : ||

QY 80 RCDSTEGTVV 89
: : || : :
Db 70 KWNDTLKQIV 79

RESULT 6
Q90B02 PRELIMINARY; PRT; 87 AA.
AC Q90B02;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PF64-49;
RA Collins K., Quinones-Mateu M., Wu M., Luze H., Johnson J., Hirsch C.,
RA Toossi Z., Arts E.;
RT "A TB-mediated site Specific Increase in HIV-1 Quasispecies Diversity
RT May Lead to the Observed Greater HIV-1 Heterogeneity in Patients Co-
RT Infected with Tuberculosis.;"
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF383521; AAK98372.1; -.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 87
SQ SEQUENCE 87 AA; 9930 MW; 1D246B516BDD1F2C CRC64;

Query Match 10.7%; Score 56; DB 15; Length 87;
Best Local Similarity 32.7%; Pred. No. 51;
Matches 17; Conservative 9; Mismatches 14; Indels 12; Gaps 4;

QY 45 NDGLTMNYCRNPDADT-----GP----WCFTTD--PSIRWEYCNLRCSDTIE 85
||: : : : ||: : : : ||: : : : ||: : : : ||
Db 16 NESVQIN-CTRPNNNRKSVHIGPGQAFYATGDIIGKIRWAYCNYSRAAWNE 66

RESULT 7
Q8H444 PRELIMINARY; PRT; 87 AA.
AC Q8H444;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE P0407H12.22 protein.
GN P0407H12.22.

```

```

OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
RT clone: P0407H12.;"
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP004303; BAC21447.1; -.
SQ SEQUENCE 87 AA; 10093 MW; 5BC46C62C68F3BF CRC64;

Query Match 10.6%; Score 55.5; DB 10; Length 87;
Best Local Similarity 30.5%; Pred. No. 58;
Matches 18; Conservative 3; Mismatches 29; Indels 9; Gaps 3;

QY 35 RHQRTPEN-YPNDGLTMNYCRNPDADTGPWCFTTDSIRWEYCNLT---RCDSTEGTVV 89
||| : : : : ||: : : : ||: : : : ||: : : : ||
Db 12 RHQAWPATAMEDDGRKDDAAAPGDSHRW-----PQIRWRPCQRPFRHCSTPPGRAV 65
||| : : : : ||: : : : ||: : : : ||: : : : ||

RESULT 8
Q8FTS2 PRELIMINARY; PRT; 85 AA.
AC Q8FTS2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN CEL489.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA Kawarabayasi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA Ikeo K., Suzuki M., Mashima J., Itch T., Yamagishi A., Nishio Y.,
RA Usuda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314.;"
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP005219; BAC18299.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 85 AA; 9868 MW; AA74412387441A4A CRC64;

Query Match 10.2%; Score 53.5; DB 16; Length 85;
Best Local Similarity 27.0%; Pred. No. 97;
Matches 17; Conservative 6; Mismatches 19; Indels 21; Gaps 4;

QY 28 WSSMTPH-----RHQRTPE-----NYPNDGLTMNYCRNPDADTGPWCFTTDSIRWE 74
||| : : : : ||: : : : ||: : : : ||: : : : ||
Db 10 WGSPLPEILGGARSYLITEGSPADGLENSKIRVNSCHR-----WCFKREISIR-D 61
||: : : : ||: : : : ||: : : : ||: : : : ||

QY 75 YCN 77
||:
Db 62 YCS 64

RESULT 9
Q9YL18 PRELIMINARY; PRT; 82 AA.
AC Q9YL18;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

```

OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.	
OX	NCBI_TaxID=11676;	
RN	[1]	
RC	SEQUENCE FROM N.A.	
RD	STRAIN=COL692;	
RE	MEDLINE=99398930; PubMed=10461834;	
RX	Navas M.C., Letourneur E., Gomas E., Boshell J., Saragosti S.;	
RY	"Analysis of the V3 loop sequences from 12 HIV type-1 infected	
RT	patients from Colombia, South America.;"	
RL	AIDS Res. Hum. Retroviruses 15:1141-1144(1999).	
RR	EMBL; Y10364; CAA71393.1; -.	
RS	InterPro; IPR000777; GP120.	
DR	Pfam; PF00516; GP120; 1.	
DK	AIDS; Coat protein; Glycoprotein.	
KW	NON_TER 1	
FT	NON_TER 88	
FT	SEQUENCE 88 AA; 9884 MW; 88ADA051D4AD1B69 CRC64;	
SQ		
	Query Match 10.08; Score 52.5; DB 15; Length 88;	
	Best Local Similarity 32.68; Pred. No. 1.3e+02;	
	Matches 15; Conservative 8; Mismatches 12; Indels 11; Gaps	
Qy	45 NDSLTMNYCRNPDAAT-----GFWCFTTD--PSIRWEYCNLFR 80	
	! : : ! ! ! : ! ! : !	
Db	13 NESVEIN-CTRPNNNTTRSHVPGKALYTTDITGDIREAYCNISR 57	
	! : : !	
RESULT 12		
Q9ID03	PRELIMINARY; PRT; 89 AA.	
ID	Q9ID03	
AC	Q9ID03;	
DT	01-OCT-2000 (TrEMBLrel. 15, Created)	
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)	
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)	
DE	Envelope protein (Fragment).	
GN	ENV OR GP120.	
OS	Human immunodeficiency virus 1.	
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.	
OX	NCBI_TaxID=11676;	
RN	[1]	
RC	SEQUENCE FROM N.A.	
RD	STRAIN=COL694;	
RE	MEDLINE=99398930; PubMed=10461834;	
RX	Navas M.C., Letourneur E., Gomas E., Boshell J., Saragosti S.;	
RY	"Analysis of the V3 loop sequences from 12 HIV type-1 infected	
RT	patients from Colombia, South America.;"	
RL	AIDS Res. Hum. Retroviruses 15:1141-1144(1999).	
RR	EMBL; Y10366; CAA71395.1; -.	
RS	InterPro; IPR000777; GP120.	
DR	Pfam; PF00516; GP120; 1.	
DK	AIDS; Coat protein; Glycoprotein.	
KW	NON_TER 1	
FT	NON_TER 89	
FT	SEQUENCE 89 AA; 9858 MW; 688C7D9A96FFE10F CRC64;	
SQ		
	Query Match 10.08; Score 52.5; DB 15; Length 89;	
	Best Local Similarity 31.68; Pred. No. 1.3e+02;	
	Matches 18; Conservative 10; Mismatches 16; Indels 13; Gaps	
Qy	45 NDSLTMNYCRNPDAAT-----GFWCFTTD--PSIRWEYCNL--TRCSDTEGTVV 89	
	! : : ! ! ! : ! ! : !	
Db	13 NESWIN-CTRPNNNTTRQGVHMGPRALYTKIVGGIRQAYCNVSRTRWNTLQIV 88	
	! : : !	
RESULT 13		
Q9MZf7	PRELIMINARY; PRT; 39 AA.	
ID	Q9MZf7	
AC	Q9MZf7;	
DT	01-OCT-2000 (TrEMBLrel. 15, Created)	
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)	
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)	
DE	Tissue plasminogen activator (Fragment).	
GN	PLAT.	

Search completed: September 29, 2003, 08:35:46
Job time : 97 secs

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	508	100.0	86	22	AAV72947		Human apolipoprotein
2	413	81.3	85	21	AAB08413		Amino acid sequenc
3	405	79.7	79	18	AAW19256		Human plasminogen
4	274.5	54.0	79	17	AAW07536		Rhesus kringle 1.
5	268.5	50.9	79	17	AAW07548		Bovine kringle 1.
6	236.5	50.9	82	21	AAV77719		Human plasminogen
7	236.5	50.9	84	10	AAP93230		Plasminogen kringle
8	236.5	50.9	84	21	AAB01906		Human plasminogen
9	236.5	50.5	79	17	AAB01906		human kringle 1.
					AAW07545		

SUMMARIES

DR N-PSDB; AAD03258.

XX PT Novel angiogenesis inhibitor, LK68 for treating angiogenesis-mediated
PT diseases, e.g. cancer and rheumatoid arthritis, has human
PT apolipoprotein(a) kringle domains IV36, IV37 and V38 amino acid
PT sequence -
XX PS Claim 3; Page 47; 50pp; English.
XX CC The present sequence is human LK8 protein which contains the amino
CC acid sequence of human apolipoprotein(a) kringle domain V38
CC (KV38). The human apolipoprotein(a) kringle domains IV36 (LK6
CC protein), IV37 (LK7 protein) and V38 (LK8 protein) together form the
CC angiogenesis inhibitor, LK68 protein. LK68, LK6, LK7 and LK8 are
CC inhibitors of endothelial cell proliferation, cell migration and
CC normal development of capillaries in the chick embryo chorioallantoic
CC membrane (CAM). LK68 protein, its single kringle or their functional
CC equivalents, are useful for treating angiogenesis-mediated diseases,
CC such as cancer, rheumatoid arthritis, psoriasis or ocular angiogenic
CC disease in animals or humans. LK68 is useful as an anticancer agent and
CC also for inhibiting primary tumour growth.
XX SQ Sequence 86 AA;
Query Match 100.0%; Score 508; DB 22; Length 86;
Best Local Similarity 100.0%; Pred. No. 1e-42;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDDCMFGNGKGYRGKATVTGTPCQEWAAQEPHRSHTFIPGKNWAGLEKNYCRNPDGD 60
Db 1 EDDCMFGNGKGYRGKATVTGTPCQEWAAQEPHRSHTFIPGKNWAGLEKNYCRNPDGD 60
QY 61 INSPWCYTMNPKLFYCDIPLCASS 86
Db 61 INSPWCYTMNPKLFYCDIPLCASS 86

RESULT 2
AAB08413
ID AAB08413 standard; Protein; 85 AA.
XX AC AAB08413;
XX DT 20-DEC-2000 (first entry)
XX DE Amino acid sequence of kringle 5 of human plasminogen.
XX KW Angiogenesis inhibitor; KED; kringle protein; angiogenesis; tPA;
KW tissue plasminogen activator; tumour; atherosclerosis; arthritis;
KW retinopathy; bronchial vascular congestion; inflammatory bowel disease;
KW adult respiratory distress syndrome; Castlemans disease; psoriasis;
KW hepatitis; aneurysm; renal disease; haemangioma.
XX OS Homo sapiens.
XX WO2000049871-A1.
XX PI 31-AUG-2000.
XX PF 24-FEB-2000; 2000WO-US04798.
XX PR 24-FEB-1999; 99US-0121341.
XX PR 25-FEB-1999; 99US-0121633.
XX PR 18-NOV-1999; 99US-0166176.
XX (FORD-) FORD HEALTH SYSTEM HENRY.
XX Dou D, Chopp M, Wang L, Mikkelsen T;
XX WPI; 2000-572016/53.
XX Use of kringle protein and kringle derived from plasminogen and
PT composition comprising kringle proteins for treating tumor and
PT atherosclerosis, arthritis and retinopathy

XX PS Disclosure; Fig 6; 163pp; English.
XX CC The specification describes a human polypeptide which is a potent
CC angiogenesis inhibitor, and is designated KED. KED polypeptide, kringle
CC proteins, or a kringle derived from human tissue plasminogen activator
CC (tPA) protein are used to inhibit angiogenesis. Kringle protein, KED
CC or tPA are useful for treating tumours, as well as atherosclerosis,
CC arthritis, retinopathy and other similar diseases. KED is also useful
CC for the treatment of diseases such as bronchial vascular congestion,
CC inflammatory bowel disease, adult respiratory distress syndrome,
CC Castlemans disease, psoriasis, hepatitis, aneurysm, renal disease
CC and haemangioma. The present sequence represents kringle 5 of human
CC plasminogen, which is used in the course of the invention.
XX SQ Sequence 85 AA;
Query Match 81.3%; Score 413; DB 21; Length 85;
Best Local Similarity 80.0%; Pred. No. 2.3e-33;
Matches 68; Conservative 7; Mismatches 10; Indels 0; Gaps 0;
QY 1 EDDCMFGNGKGYRGKATVTGTPCQEWAAQEPHRSHTFIPGKNWAGLEKNYCRNPDGD 60
Db 1 EDDCMFGNGKGYRGKATVTGTPCQEWAAQEPHRSHTFIPGKNWAGLEKNYCRNPDGD 60
QY 61 INSPWCYTMNPKLFYCDIPLCASS 85
Db 61 VGGPWCTTNDKRLDYCDVPOCAA 85
RESULT 3
AAW19256
ID AAW19256 standard; Peptide; 79 AA.
XX AC AAW19256;
XX DT 27-FEB-1998 (first entry)
XX DE Human plasminogen Kringle 5 fragment.
XX KW Plasminogen; Kringle 5; cell proliferation inhibitor; angiogenesis;
KW diagnosis; therapeutic.
XX OS Homo sapiens.
XX PN WO9723500-A1.
XX PD 03-JUL-1997.
XX PF 13-DEC-1996; 96WO-US20447.
XX PR 12-DEC-1996; 96US-0763528.
XX PR 13-DEC-1995; 95US-0008519.
XX (CHIL-) CHILDRENS MEDICAL CENT.
XX PI Cao Y, Folkman MJ;
XX WPI; 1997-350965/32.
XX Plasminogen Kringle 5 peptide - which inhibits endothelial cell
PT proliferation, useful to treat angiogenesis mediated diseases and in
PT detection and diagnosis
XX Claim 1; Page 8; 51pp; English.
XX This sequence is an isolated fragment of the Kringle 5 peptide
CC corresponding to amino acid 462 of the human plasminogen protein which
CC can be used in a novel method to inhibit endothelial cell proliferation
CC activity. The protein can be used to treat angiogenesis mediated
CC diseases, e.g. haemangioma, solid tumours, leukaemia, metastasis,
CC telangiectasia, psoriasis, scleroderma, pyogenic granuloma, myocardial
CC angiogenesis, plaque neovascularisation, coronary or cerebral

CC acids 6-84 of bovine angiotensin.
XX Sequence 79 AA;
SQ Query Match 50.9%; Score 258.5; DB 17; Length 79;
Best Local Similarity 57.5%; Pred. No. 3.5e-18;
Matches 46; Conservative 7; Mismatches 26; Indels 1; Gaps 1;
QY 4 CMFGNGKGYRGKATVTGTPCQEWAAQEPHRSHTFPGTNKAGLEKNYCRNPDDING 63
DB 1 CKTNGQYRGTAETKSGVTCKWSATSPH-VKFPSPKFPPLAGLEENYCRNPDDENG 59
QY 64 PWCYTNNPKLFDYCDIPLC 83
DB 60 PWCYTTPDKRYDYCDIPEC 79
RESULT 6
ID AAY77719 standard; protein; 82 AA.
XX AAY77719;
AC AAY77719;
XX 12-MAY-2000 (first entry)
XX Human plasminogen kringle 1 sequence.
XX Kringle 1; plasminogen; anti-angiogenic; angiogenesis; angiotensin;
KW cytostatic; antiarthritic; antirheumatic; antidiabetic; ophthalmological;
KW immunosuppressant; vasotropic; vulnery; antiarteriosclerotic; human;
KW dermatological; cancer; tumour; birth control; vascularization.
XX Homo sapiens.
OS
XX
XX
XX Key Location/Qualifiers
FT Region 36..73
FT /note= "specifically claimed fragment (AAY77722)"
XX WO200003726-A1.
XX 27-JAN-2000.
XX 07-JUL-1999; 99WO-US15271.
XX 14-JUL-1998; 98US-0092831.
XX (BRIM) BRISTOL-MYERS SQUIBBS CO.
XX Ji R, Trail PA;
XX WPI; 2000-171200/15.
XX Novel lysine binding fragments angiotensin used as antiangiogenic
PT agents in the treatment of cancer, diabetic retinopathy, rheumatoid
PT arthritis, psoriasis, atherosclerotic plaque formation, and other
PT angiogenesis diseases -
XX Disclosure; Fig 2; 30pp; English.
XX The invention provides fragments of kringle 1, 2 or 4 of human
CC plasminogen that contain a lysine binding site and have anti-angiogenic
CC activity. The peptides of the invention function as antiangiogenic
CC agents, for the treatment process and diseases involving angiogenesis.
CC Such diseases include cancers such as solid tumours, blood born tumours
CC such as leukemias, tumours metastasis, benign tumours such as
CC hemangiomas, acoustic acouromas, neurofibromas, trachomas and pyogenic
CC granulomas, rheumatoid arthritis, ocular angiogenic diseases such as
CC diabetic retinopathy, retinopathy of prematurity, macular degeneration,
CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasia,
CC rheiosis, Osler-Webber syndrome, myocardial angiogenesis, plaque
CC neovascularization, telangiectasia, hemophilic joints, angiofibroma,
CC and wound granulation. The fragments are also useful in treatment of
CC disease of excessive or abnormal stimulation of endothelial cells. These

CC diseases include intestinal adhesions, atherosclerosis, scleroderma, and
CC hypertrophic scars. The fragments can be used as birth control agents by
CC preventing vascularization required for embryo transplantation. The
CC present sequence represents the kringle 1 sequence of human plasminogen
CC (angiotensin).
XX
SQ Sequence 82 AA;
Query Match 50.9%; Score 258.5; DB 21; Length 82;
Best Local Similarity 54.3%; Pred. No. 3.7e-18;
Matches 44; Conservative 9; Mismatches 27; Indels 1; Gaps 1;
QY 3 DCMFGNGKGYRGKATVTGTPCQEWAAQEPHRSHTFPGTNKAGLEKNYCRNPDDING 62
DB 2 ECKTNGKNGYRGTMSTKNGITCKKWSSTSPHR-PRESPATHPSEGLEENYCRNPDDNQ 60
QY 63 GPWCYTNNPKLFDYCDIPLC 83
DB 61 GPWCYTTPDKRYDYCDILEC 81
RESULT 7
ID AAP93230 standard; peptide; 84 AA.
XX AAP93230;
AC AAP93230;
XX 25-MAR-2003 (updated)
DT 03-APR-1990 (first entry)
XX Plasminogen kringle 1 domain (residues 79-162).
DE Plasminogen; activator; t-PA; fibrin; kringle domain;
KW scu-PA;
XX Homo sapiens.
OS
XX WO8910401-A.
XX 02-NOV-1989.
XX 23-MAR-1989; 89WO-US01255.
XX 22-APR-1988; 88US-0184823.
XX (COLB) COLLABORATIVE RES INC.
XX Mao JI, Abercrombie DM;
PI WPI; 1989-339965/46.
DR N-PSDB; AAN92237.
XX Modified plasminogen activator - having greater fibrin
PT selectivity and circulating half-life.
XX Disclosure; fig 3A; 80pp; English.
XX The sequence is encoded by a synthetic oligonucleotide and is the kringle
CC 1 domain of plasminogen. The DNA is used to construct a modified
CC plasminogen activator (mPA) of greater fibrin selectivity than normal PA.
CC The mPA is used in the treatment of pulmonary embolism, thrombosis,
CC myocardial infarct and strokes.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 84 AA;
Query Match 50.9%; Score 258.5; DB 10; Length 84;
Best Local Similarity 54.3%; Pred. No. 3.8e-18;
Matches 44; Conservative 9; Mismatches 27; Indels 1; Gaps 1;
QY 3 DCMFGNGKGYRGKATVTGTPCQEWAAQEPHRSHTFPGTNKAGLEKNYCRNPDDING 62
DB 5 ECKTNGKNGYRGTMSTKNGITCKKWSSTSPHR-PRESPATHPSEGLEENYCRNPDDNQ 63

QY 63 GPWCYTMNPKRLFDYCDIPLC 83
 Db 64 GPWCYITDPEKRYDYCDILEC 84

RESULT 8
 AAB01906
 ID AAB01906 standard; Protein; 84 AA.

AC AAB01906;
 XX
 DT 18-SEP-2000 (first entry)
 XX
 DE Human plasminogen kringle 1 (Tyr80-Glu163).

XX Plasminogen; human; kringle domain; endothelial cell proliferation;
 KW angiogenesis; antiproliferative; antiarteriosclerotic; cytostatic;
 KW antiporatic; antiinflammatory; antitumor; antirheumatic; antiarthritic;
 KW antiangiogenic; cancer; tumour; autoimmune disease.

XX Homo sapiens.

XX US6057122-A.

PN 02-MAY-2000.

XX 05-MAY-1997; 97US-0851350.

XX 03-MAY-1996; 96US-0643219.

PR 03-APR-1997; 97US-0832087.

XX (ABBO) ABBOTT LAB.

XX Davidson DJ;

XX WPI; 2000-349573/30.

XX Preparation of Kringle five peptide fragment for treating various
 PT disorders such as angiogenic, ocular, skin diseases and cancer,
 PT involves mixing mammalian plasminogen and elastase followed by
 PT incubation and isolation -

PS Example 17; Page -: 48pp; English.

XX The invention relates to a method of preparing plasminogen kringle 5
 CC peptide fragments. The method comprises mixing mammalian plasminogen and
 CC elastase in the ratio 1:100-1:300, followed by incubating and isolating
 CC the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
 CC endothelial cell proliferation and migration. The peptides are useful
 CC for treating angiogenic diseases, primary and metastatic solid tumours
 CC and carcinomas of various organs such as breast, genital tract,
 CC endocrine glands, skin, tumours of the brain and eyes and solid tumours
 CC arising from haematopoietic malignancies such as leukaemias and
 CC lymphomas. They are also used for the prophylaxis of various autoimmune
 CC diseases (e.g., rheumatoid arthritis), ocular diseases, skin diseases
 CC (e.g., psoriasis), blood vessel diseases (e.g., haemangiomas, Osler-Webber
 CC syndrome), diseases caused by excessive or abnormal stimulation of
 CC endothelial cells (e.g., Crohn's disease, atherosclerosis), diseases
 CC which have angiogenesis as a pathologic consequence (e.g., cat scratch
 CC disease and ulcers). The peptides are also useful as a birth control
 CC agent which inhibits ovulation and establishment of the placenta.
 CC Sequences AAB01906-B01919 represent fragments of human plasminogen used
 CC in an exemplification of the invention.

CC Note: This sequence is not shown in the specification, but is derived
 CC from the full length human plasminogen sequence (AAB01887) shown in
 CC figure 1.

XX Sequence 84 AA;

Query Match 50.9%; Score 258.5; DB 21; Length 84;
 Best Local Similarity 54.3%; Pred. No. 3.8e-18;
 Matches 44; Conservative 9; Mismatches 27; Indels 1; Gaps 1;

QY 3 DCMFGNGKGYRGKATVTGTGTCQEWAAQEPHRSHTFIPGTNKGAGLEKNYCRNPDGDN 62
 Db 4 ECKTGNKNGYRGKATVTGTGTCQEWAAQEPHRSHTFIPGTNKGAGLEKNYCRNPDGDN 62

QY 63 GPWCYTMNPKRLFDYCDIPLC 83

Db 63 GPWCYITDPEKRYDYCDILEC 83

RESULT 9
 AAW07545
 ID AAW07545 standard; protein; 79 AA.

XX AAW07545;

XX 21-JUN-1997 (first entry)

XX human kringle 1.

XX angiotatin; plasminogen; kringle; angiogenesis; cancer; arthritis;
 KW macular degeneration; diabetic retinopathy.

XX Homo sapiens.

XX WO9635774-A2.

XX 14-NOV-1996.

XX 26-APR-1996; 96WO-US05856.

XX 08-MAR-1996; 96US-0612788.

XX 26-APR-1995; 95US-0429743.

XX 22-FEB-1996; 96US-0605598.

XX (CHIL-) CHILDRENS MEDICAL CENT.

XX Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;

XX WPI; 1996-518662/51.

XX Use of angiotatin fragments or aggregates - for inhibiting
 PT endothelial cell proliferation and treating angiogenesis-mediated
 PT diseases, e.g. cancer, arthritis or diabetic retinopathy

XX Claim 4; Page 104; 203pp; English.

XX The invention relates to new methods and compositions for
 CC inhibiting endothelial cell proliferation, using as active component
 CC an angiotatin fragment, a combination of angiotatin fragments, or
 CC aggregate angiotatin. The fragment is preferably derived from murine,
 CC human, Rhesus, porcine or bovine plasminogen and is a kringle 1,
 CC kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle
 CC 1-4 or kringle 1-4BKLS protein. The aggregate angiotatin has a Mol. wt.
 CC of 45-65 kD and is derived from a plasminogen fragment beginning at
 CC approximately amino acid number 98 of murine, human, Rhesus, porcine or
 CC bovine plasminogen. The active component can be used for treating
 CC angiogenesis-mediated diseases such as cancer, arthritis, macular
 CC degeneration and diabetic retinopathy. It can also be used to develop
 CC antibodies for use in diagnosis, detection and therapy.
 CC The present sequence, human kringle 1, is a specific angiotatin
 CC fragment which can be used in the invention, and represents amino
 CC acids 6-84 of human angiotatin.

XX Sequence 79 AA;

Query Match 50.5%; Score 256.5; DB 17; Length 79;

Best Local Similarity 55.0%; Pred. No. 5.6e-18;

Matches 44; Conservative 8; Mismatches 27; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATVTGTGTCQEWAAQEPHRSHTFIPGTNKGAGLEKNYCRNPDGDN 63

Db 1 CKTGNKNGYRGKATVTGTGTCQEWAAQEPHRSHTFIPGTNKGAGLEKNYCRNPDGDN 59

```

QY      64  PWCYTMNPKLFYCDIPLC 83
DQ      60  PWCYTDPEKRYDYCDILEC 79

RESULT 10
AAP90598
ID  AAP90598 standard; protein; 81 AA.
XX  AAP90598;
XX
DT  25-MAR-2003 (updated)
DT  31-OCT-2002 (updated)
DT  13-JUN-1990 (first entry)
XX
DE  Synthetic N-terminal protein of truncated human tissue plasminogen
DE  activator (t-PA).
XX
KW  t-PA; tissue plasminogen activator; thrombolytic;
KW  fibrinolytic; fibrin; ds;
XX
OS  Synthetic.
XX
PN  WO8900191-A.
XX
PD  12-JAN-1989.
XX
PF  06-JUL-1988; 88WO-US02263.
XX
PR  06-JUL-1987; 87US-0070012.
PR  20-APR-1988; 88US-0184121.
XX
PA  (GEMY ) GENETICS INST INC.
XX
PI  Larsen GR, Ahern T, Langersafe P;
DR  WPT; 1989-039648/05.
DR  N-PSDB; AAN91194.
XX
PT  Tissue plasminogen activator variants -
PT  having N-terminal domain replaced with first kringle region
PT  domain of plasminogen.
XX
PS  Claim 1; Page 32; 36pp; English.
XX
CC  Inserted in front of an N-terminal 1-82 AA truncated t-PA protein, new
CC  product has an increased affinity to fibrin, decreased reactivity with
CC  inhibitors and increased thrombolytic and fibrinolytic activity.
CC  (Updated on 31-OCT-2002 to add missing OS field.)
CC  (Updated on 25-MAR-2003 to correct PA field.)
CC  (Updated on 25-MAR-2003 to correct PI field.)
CC  (Updated on 25-MAR-2003 to correct DR field.)
XX
SQ  Sequence 81 AA;

Query Match      50.0%; Score 254; DB 10; Length 81;
Best Local Similarity 53.1%; Pred. No. 1e-17;
Matches 43; Conservative 9; Mismatches 27; Indels 2; Gaps 1;

QY      3  DCMFNGKGYRGKATVTGTPQEWAAQEPHRHSTFTPGTNKAGLEKNYCRNPDGIN 62
DQ      2  ECKTGNKGYRTMSKYNKATGCKWSTSP--HPRFSPATHPSEGLEENYCRNPDNPQ 59

QY      63  GPWCYTMNPKLFYCDIPLC 83
DQ      60  GPWCYTDPEKRYDYCDILEC 80

RESULT 11
AAB08410
ID  AAB08410 standard; Protein; 84 AA.
XX

```

```

AC      AAB08410;
XX
DT  20-DEC-2000 (first entry)
XX
DE  Amino acid sequence of kringle 1 of human plasminogen.
XX
KW  Angiogenesis inhibitor; KED; kringle protein; angiogenesis; tPA;
KW  tissue plasminogen activator; tumour; atherosclerosis; arthritis;
KW  retinopathy; bronchial vascular congestion; inflammatory bowel disease;
KW  adult respiratory distress syndrome; Castlemans disease; psoriasis;
KW  hepatitis; aneurysm; renal disease; haemangioma.
XX
OS  Homo sapiens.
XX
PN  WO2000049871-A1.
XX
PD  31-AUG-2000.
XX
PF  24-FEB-2000; 2000WO-US04798.
XX
PR  24-FEB-1999; 99US-0121341.
PR  26-FEB-1999; 99US-0121633.
PR  18-NOV-1999; 99US-0166176.
XX
PA  (FORD-) FORD HEALTH SYSTEM HENRY.
XX
PI  Dou D, Chopp M, Wang L, Mikkelsen T;
XX
DR  WPI; 2000-572016/53.
XX
PT  Use of kringle protein and kringle derived from plasminogen and
PT  composition comprising kringle proteins for treating tumor and
PT  atherosclerosis, arthritis and retinopathy
XX
PS  Disclosure; Fig 6; 163pp; English.
XX
CC  The specification describes a human polypeptide which is a potent
CC  angiogenesis inhibitor, and is designated KED. KED polypeptide, kringle
CC  proteins, or a kringle derived from human tissue plasminogen activator
CC  (tPA) protein are used to inhibit angiogenesis. Kringle protein, KED
CC  or tPA are useful for treating tumours, as well as atherosclerosis,
CC  arthritis, retinopathy and other similar diseases. KED is also useful
CC  for the treatment of diseases such as bronchial vascular congestion,
CC  inflammatory bowel disease, psoriasis, hepatitis, aneurysm, renal disease
CC  and haemangioma. The present sequence represents kringle 1 of human
CC  plasminogen, which is used in the course of the invention.
XX
SQ  Sequence 84 AA;

Query Match      49.7%; Score 252.5; DB 21; Length 84;
Best Local Similarity 53.1%; Pred. No. 1.5e-17;
Matches 43; Conservative 8; Mismatches 29; Indels 1; Gaps 1;

QY      3  DCMFNGKGYRGKATVTGTPQEWAAQEPHRHSTFTPGTNKAGLEKNYCRNPDGIN 62
DQ      3  ECKTGNKGYRTMCKTKNGTCKWSTSPHR-PFSPATHDSEGLEENYCRNPDNPQ 61

QY      63  GPWCYTMNPKLFYCDIPLC 83
DQ      62  GPWCYTDPEKRYDYCDILEC 82

RESULT 12
AAP81360
ID  AAP81360 standard; protein; 79 AA.
XX
AC  AAP81360;
XX
DT  25-MAR-2003 (updated)
DT  31-OCT-2002 (updated)
DT  06-OCT-1990 (first entry)
XX

```

```

DE KI domain of plasminogen.
XX
KW Tissue plasminogen activator; analogue; K1 domain.
XX
OS Homo sapiens.
XX
PN EF293934-A.
XX
XX 07-DEC-1988.
XX
XX 03-JUN-1988; 88EP-0108949.
XX
XX 04-JUN-1987; 87US-0058217.
XX
XX (ZYMO ) ZYMOGENETICS INC.
XX
XX (NOVO ) NOVO IND AS.
XX
XX (EISA ) EISA CO LTD.
XX
XX Mulvihill ER, Nexo BA, Yoshitake S, Ikeda Y, Suzuki S;
XX Hashimoto A, Yuzuriha T;
XX
XX WPI; 1988-347625/49.
XX
XX N-PSDB; AAN81087.
XX
XX New tissue plasminogen activator analogues with non-native K1-domain
XX having higher fibrin specificity, and new encoding DNA sequences.
XX
XX Disclosure; Page ?; ?pp; English.
XX
XX The K1 domain of plasminogen may be used to replace the K1 domain of
XX t-PA to produce an analogue with greater specificity for fibrin than
XX native t-PA.
XX
XX (Updated on 31-OCT-2002 to add missing OS field.)
XX
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
XX (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 79 AA;

Query Match 49.5%; Score 251.5; DB 9; Length 79;
Best Local Similarity 53.8%; Pred. No. 1.7e-17;
Matches 43; Conservative 9; Mismatches 27; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATVTGTGTCQEWAAQEPHRSHTFIPGTNKWAGLEKNCYCRNPDGDI 63
DB 1 CKTGDKNYRGTSKTKNGITCKQWSSSTSPHR-PRFSPATHPSEGLEENYCRNPDNDPQG 59

QY 64 PWCYTMNPKLFYDCDILPC 83
DB 60 PWCYTTDPKRYDYCDILEC 79

RESULT 13
AAR96221
ID AAR96221 standard; Protein; 79 AA.
XX
XX AAR96221;
XX
XX 25-MAR-2003 (updated)
XX 06-AUG-1996 (first entry)
XX
XX Plasminogen kringle domain K1.
XX
XX Tissue plasminogen activator; tPA; alpha2-plasmin inhibitor;
XX fibrinolytic; thrombolytic; fibrin; thrombosis; blood clotting;
XX protein engineering; kringle.
XX
XX Synthetic.
XX
XX US5504001-A.
XX
XX 02-APR-1996.
XX
XX 06-JUN-1994; 94US-0254485.

DE KI domain of plasminogen.
XX
KW Tissue plasminogen activator; analogue; K1 domain.
XX
OS Homo sapiens.
XX
PN EF293934-A.
XX
XX 07-DEC-1988.
XX
XX 03-JUN-1988; 88EP-0108949.
XX
XX 04-JUN-1987; 87US-0058217.
XX
XX (ZYMO ) ZYMOGENETICS INC.
XX
XX (NOVO ) NOVO IND AS.
XX
XX (EISA ) EISA CO LTD.
XX
XX Mulvihill ER, Nexo BA, Yoshitake S, Ikeda Y, Suzuki S;
XX Hashimoto A, Yuzuriha T;
XX
XX WPI; 1988-347625/49.
XX
XX N-PSDB; AAN81087.
XX
XX New tissue plasminogen activator analogues with non-native K1-domain
XX having higher fibrin specificity, and new encoding DNA sequences.
XX
XX Disclosure; Page ?; ?pp; English.
XX
XX The K1 domain of plasminogen may be used to replace the K1 domain of
XX t-PA to produce an analogue with greater specificity for fibrin than
XX native t-PA.
XX
XX (Updated on 31-OCT-2002 to add missing OS field.)
XX
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
XX (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 79 AA;

Query Match 49.5%; Score 251.5; DB 17; Length 79;
Best Local Similarity 53.8%; Pred. No. 1.7e-17;
Matches 43; Conservative 9; Mismatches 27; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATVTGTGTCQEWAAQEPHRSHTFIPGTNKWAGLEKNCYCRNPDGDI 63
DB 1 CKTGDKNYRGTSKTKNGITCKQWSSSTSPHR-PRFSPATHPSEGLEENYCRNPDNDPQG 59

QY 64 PWCYTMNPKLFYDCDILPC 83
DB 60 PWCYTTDPKRYDYCDILEC 79

RESULT 14
AAR94401
ID AAR94401 standard; Protein; 78 AA.
XX
XX AAR94401;
XX
XX 01-JUL-2002 (updated)
XX 18-JUN-1990 (first entry)
XX
XX Sequence encoding K1 domain of plasminogen.
XX
XX Tissue plasminogen activator; fibrinolytic; t-PA; plasminogen.
XX
XX Homo sapiens.
XX
XX A08817430-A.
XX
XX 08-DEC-1988.
XX
XX 06-JUN-1988; 88AU-0017430.
XX
XX 04-JUN-1987; 87JP-0058061.
XX 02-MAR-1988; 88JP-0062847.
XX
XX (NOVO ) NOVO INDUSTI A/S.
XX (ZYMO ) ZYMOGENETICS INC.
XX (EISA ) EISA CO LTD.
XX
XX Mulvihill ER, Ohara PJ, Nexo BA, Yoshitake S, Ikeda Y, Suzuki S;

```

PI Hashimoto A, Yuzuhira T;
 XX WPI; 1989-039773/06.
 DR N-PSDB; AAN91218.
 XX
 PT Tissue plasminogen activator analogues -
 PT having at least one cysteine residue in the growth factor domain
 PT replaced with another amino acid.
 XX
 PS Example 9; Fig 11; 95pp; English.
 XX
 CC Amino acid sequence of kringle domain of plasminogen, incorporated into
 CC plasmid PK1. Sequence is derived from 11 synthetic oligonucleotides.
 CC Alternatively the sequence is described as encoding Asn at position 96.
 CC (Updated on 01-JUL-2002 to add missing PI field.)
 XX
 SQ Sequence 78 AA;
 Query Match 48.9%; Score 248.5; DB 10; Length 78;
 Best Local Similarity 54.5%; Pred. No. 3.4e-17;
 Matches 42; Conservative 9; Mismatches 25; Indels 1; Gaps 1;
 QY 4 CMFGNGKGYRGKATVTGTGTCQEWAAQEPHRSHTFIPGTNKWAGLEKNCYCRNPDG 63
 Db 1 CTGDKGNKRGTSKTKGTCQKWSSTSPH-PRFSPATHSEGLENCYCRNPDNDPQG 59
 QY 64 PWCYTMDPKRFDYCDI 80
 Db 60 PWCYTMDPKRFDYCDI 76
 RESULT 15
 AAW07547
 ID AAW07547 standard; protein; 79 AA.
 AC AAW07547;
 XX
 XX 21-JUN-1997 (first entry)
 XX Porcine kringle 1.
 XX
 DE angiotatin; plasminogen; kringle; angiogenesis; cancer; arthritis;
 KW macular degeneration; diabetic retinopathy.
 XX
 XX Sus scrofa.
 XX WO9635774-A2.
 XX
 XX 14-NOV-1996.
 XX
 XX 26-APR-1996; 96WO-0505856.
 XX
 XX 08-MAR-1996; 96US-0612788.
 XX
 XX 26-APR-1995; 95US-0429743.
 XX
 XX 22-FEB-1996; 96US-0605598.
 XX
 XX (CHIL-) CHILDRENS MEDICAL CENT.
 XX
 XX Cao Y, Folkman MJ, Lin J, O'Reilly MS, Sim KL;
 XX WPI; 1996-518662/51.
 XX
 PT Use of angiotatin fragments or aggregates - for inhibiting
 PT endothelial cell proliferation and treating angiogenesis-mediated
 PT diseases, e.g. cancer, arthritis or diabetic retinopathy
 XX
 XX Claim 4; Page 105; 203pp; English.
 XX
 CC The invention relates to new methods and compositions for
 CC inhibiting endothelial cell proliferation, using as active component
 CC an angiotatin fragment, a combination of angiotatin fragments, or
 CC aggregate angiotatin. The fragment is preferably derived from murine,
 CC human, Rhesus, porcine or bovine plasminogen and is a kringle 1,

CC kringle 2, kringle 3, kringle 2-3, kringle 1-3, kringle 1-2, kringle
 CC 1-4 or kringle 1-4BKLS protein. The aggregate angiotatin has a Mol. Wt.
 CC of 45-65 kD and is derived from a plasminogen fragment beginning at
 CC approximately amino acid number 98 of murine, human, Rhesus, porcine or
 CC bovine plasminogen. The active component can be used for treating
 CC angiogenesis-mediated diseases such as cancer, arthritis, macular
 CC degeneration and diabetic retinopathy. It can also be used to develop
 CC antibodies for use in diagnosis, detection and therapy.
 CC The present sequence, porcine kringle 1, is a specific angiotatin
 CC fragment which can be used in the invention, and represents amino
 CC acids 6-84 of porcine angiotatin.
 XX
 SQ Sequence 79 AA;

Query Match 48.5%; Score 246.5; DB 17; Length 79;
 Best Local Similarity 53.8%; Pred. No. 5.4e-17;
 Matches 43; Conservative 7; Mismatches 29; Indels 1; Gaps 1;
 QY 4 CMFGNGKGYRGKATVTGTGTCQEWAAQEPHRSHTFIPGTNKWAGLEKNCYCRNPDG 63
 Db 1 CXTGNGKNYRGTTSTKTKSGVICQKWSVSPH-IPKYSPEKFFLAGLENCYCRNPDNDKG 59
 QY 64 PWCYTMDPKRFDYCDI 83
 Db 60 PWCYTMDPKRFDYCDI 79

Search completed: September 29, 2003, 08:46:58
 Job time : 83 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:45:34 ; Search time 29 Seconds
(without alignments)
125.474 Million cell updates/sec

Title: US-10-088-548-8
Perfect score: 508
Sequence: 1 BODCMFGNGKRGKATV.....YTMNPKLFDDYCDIPLASS 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 222816

Minimum DB seq length: 0
Maximum DB seq length: 86

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2.6/ptodata/1/aaa/5A_COMB.pep.*
2: /cgn2.6/ptodata/1/aaa/5B_COMB.pep.*
3: /cgn2.6/ptodata/1/aaa/5A_COMB.pep.*
4: /cgn2.6/ptodata/1/aaa/6B_COMB.pep.*
5: /cgn2.6/ptodata/1/aaa/PCTUS_COMB.pep.*
6: /cgn2.6/ptodata/1/aaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	412	81.1	80	2 US-08-763-528A-6	Sequence 6, Appl
2	405	79.7	79	2 US-08-763-528A-1	Sequence 1, Appl
3	274.5	54.0	79	2 US-08-612-788-9	Sequence 9, Appl
4	274.5	54.0	79	3 US-09-066-028-9	Sequence 9, Appl
5	274.5	54.0	79	4 US-09-335-325-9	Sequence 9, Appl
6	258.5	50.9	79	2 US-08-612-788-11	Sequence 11, Appl
7	258.5	50.9	79	3 US-09-066-028-11	Sequence 11, Appl
8	258.5	50.9	79	4 US-09-335-325-11	Sequence 11, Appl
9	258.5	50.9	82	4 US-09-348-953-1	Sequence 1, Appl
10	256.5	50.5	79	2 US-08-612-788-8	Sequence 8, Appl
11	256.5	50.5	79	3 US-08-763-528A-2	Sequence 2, Appl
12	256.5	50.5	79	3 US-09-066-028-8	Sequence 8, Appl
13	256.5	50.5	79	4 US-09-335-325-8	Sequence 8, Appl
14	246.5	48.5	79	2 US-08-612-788-10	Sequence 10, Appl
15	246.5	48.5	79	3 US-09-066-028-10	Sequence 10, Appl
16	246.5	48.5	79	4 US-09-335-325-10	Sequence 10, Appl
17	241.5	47.5	79	2 US-08-612-788-7	Sequence 7, Appl
18	241.5	47.5	79	3 US-09-066-028-7	Sequence 7, Appl
19	241.5	47.5	79	4 US-09-335-325-7	Sequence 7, Appl
20	240	47.2	78	2 US-08-612-788-22	Sequence 22, Appl
21	240	47.2	78	3 US-09-066-028-22	Sequence 22, Appl
22	240	47.2	78	4 US-09-335-325-22	Sequence 22, Appl
23	235	46.3	84	4 US-09-348-953-2	Sequence 2, Appl
24	222	43.7	78	2 US-08-612-788-17	Sequence 17, Appl
25	222	43.7	78	2 US-08-612-788-23	Sequence 23, Appl
26	222	43.7	78	2 US-08-763-528A-5	Sequence 5, Appl
27	222	43.7	78	3 US-09-066-028-17	Sequence 17, Appl

Sequence 23, Appl
Sequence 17, Appl
Sequence 23, Appl
Sequence 16, Appl
Sequence 20, Appl
Sequence 16, Appl
Sequence 20, Appl
Sequence 16, Appl
Sequence 20, Appl
Sequence 16, Appl
Sequence 20, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 14, Appl
Sequence 13, Appl
Sequence 14, Appl
Sequence 13, Appl
Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-08-763-528A-6
; Sequence 6, Application US/08763528A
; Patent No. 5854221
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Endothelial Cell Proliferation Inhibitor
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew, LLP
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/763,528A
; FILING DATE: 12-DEC-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05940-0251
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..80
; OTHER INFORMATION: /note= "Kingle 5 - Figure 3"
US-08-763-528A-6

Query Match 81.1%; Score 412; DB 2; Length 80;


```
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S
/ ZIP: 30303-1769
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: IBM PC compatible
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION NUMBER: US/09/066,028
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/612,788
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 79 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Rhesus monkey
/ IMMEDIATE SOURCE:
/ CLONE: K1
/
/ US-09-066-028-9
/
/ Query Match 54.0%; Score 274.5; DB 3; Length 79;
/ Best Local Similarity 58.8%; Pred. No. 3.9e-24;
/ Matches 47; Conservative 8; Mismatches 24; Indels 1; Gaps 1;
/
/ QY 4 CMFGNGKGYRGKKATVVTGTCQEWAAQEPHRHSTFIPGINKWAGLEKKNYCRNPDGDLNG 63
/ Db 1 CKTGNGKNYRGTSKRTGTGTCQKWSSTSPHR-PTFSPATHPSEGLEENYCRNPDNDGQG 59
/
/ QY 64 PWCYTMNPKLFYDCDIPLC 83
/ Db 60 PWCYTTDEERFYCDIPEC 79
/
/ RESULT 5
/ US-09-335-325-9
/ Sequence 9, Application US/09335325
/ Patent No. 6521439
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ O'Reilly, Micheal
/ Cao, Yihai
/ Sim, B. Kim Lee
/
/ TITLE OF INVENTION: Angiostatin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S
/ ZIP: 30303-1769
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
```

```
/
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/335,325
/ FILING DATE: 17-Jun-1999
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 79 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Rhesus monkey
/ IMMEDIATE SOURCE:
/ CLONE: K1
/
/ SEQUENCE DESCRIPTION: SEQ ID NO: 9:
/
/ US-09-335-325-9
/
/ Query Match 54.0%; Score 274.5; DB 4; Length 79;
/ Best Local Similarity 58.8%; Pred. No. 3.9e-24;
/ Matches 47; Conservative 8; Mismatches 24; Indels 1; Gaps 1;
/
/ QY 4 CMFGNGKGYRGKKATVVTGTCQEWAAQEPHRHSTFIPGINKWAGLEKKNYCRNPDGDLNG 63
/ Db 1 CKTGNGKNYRGTSKRTGTGTCQKWSSTSPHR-PTFSPATHPSEGLEENYCRNPDNDGQG 59
/
/ QY 64 PWCYTMNPKLFYDCDIPLC 83
/ Db 60 PWCYTTDEERFYCDIPEC 79
/
/ RESULT 6
/ US-08-612-788-11
/ Sequence 11, Application US/08612788
/ Patent No. 5837682
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ O'Reilly, Micheal
/ Cao, Yihai
/ APPLICANT: Sim, B. Kim Lee
/
/ TITLE OF INVENTION: Angiostatin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S
/ ZIP: 30303-1769
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: IBM PC compatible
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE:
```



```
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Bovine
; IMMEDIATE SOURCE:
; CLONE: K1
; US-08-612-788-11
;
; Query Match 50.9%; Score 258.5; DB 2; Length 79;
; Best Local Similarity 57.5%; Pred. No. 2.6e-22;
; Matches 46; Conservative 7; Mismatches 26; Indels 1; Gaps 1;
;
; QY 4 CMFGNGKGYRGKATVTGTCQEWAAQPHRHSFIPTGKNWAGLEKNCRNPDGDDING 63
; DB 1 CKTGNGQTYRGTAETKSGVTCQKWSATSPH-VPKFSEKFPFLAGLEENTCRPNPDENG 59
;
; QY 64 PWCYTMNPKLFYDCDIPLC 83
; DB 60 PWCYTTDPKRYDYCDIPEC 79
;
; RESULT 7
; US-09-066-028-11
; Sequence 11, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
```

```
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Bovine
; IMMEDIATE SOURCE:
; CLONE: K1
; US-09-066-028-11
;
; Query Match 50.9%; Score 258.5; DB 3; Length 79;
; Best Local Similarity 57.5%; Pred. No. 2.6e-22;
; Matches 46; Conservative 7; Mismatches 26; Indels 1; Gaps 1;
;
; QY 4 CMFGNGKGYRGKATVTGTCQEWAAQPHRHSFIPTGKNWAGLEKNCRNPDGDDING 63
; DB 1 CKTGNGQTYRGTAETKSGVTCQKWSATSPH-VPKFSEKFPFLAGLEENTCRPNPDENG 59
;
; QY 64 PWCYTMNPKLFYDCDIPLC 83
; DB 60 PWCYTTDPKRYDYCDIPEC 79
;
; RESULT 8
; US-09-335-325-11
; Sequence 11, Application US/09335325
; Patent No. 6521439
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
```

```
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Bovine
/ IMMEDIATE SOURCE:
/ CLONE: K1
/ SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-335-325-11

Query Match          50.9%; Score 258.5; DB 4; Length 79;
Best Local Similarity 57.5%; Pred. No. 2.6e-22;
Matches 46; Conservative 7; Mismatches 26; Indels 1; Gaps 1;

QY 4 CMFNGKGYRGKATVITGTPCQEWAAQEPHRSHTFIPGTNKWAGLEKNYCRNPDGDI 63
Db 1 CKTGNGKNGYRGTMSTKNGITCQKWSSTSPHR-PFSPATHPSEGLEENYCRNPDNDG 59
QY 64 PWCYTMNPKLFYDCDIPLC 83
Db 60 PWCITTDPRYDYCDILEC 79

RESULT 9
US-09-348-953-1
/ Sequence 1, Application US/09348953
/ Patent No. 6538103
/ GENERAL INFORMATION:
/ APPLICANT: Ji, Richard Wei-Dong
/ APPLICANT: Trill, Pamela A.
/ TITLE OF INVENTION: LYSINE BINDING FRAGMENTS OF ANGIOSTATIN
/ FILE REFERENCE: DB11sequences
/ CURRENT APPLICATION NUMBER: US/09/348,953
/ CURRENT FILING DATE: 1999-07-07
/ PRIOR APPLICATION NUMBER: 60/092,831
/ PRIOR FILING DATE: 1998-07-14
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 82
/ TYPE: PRT
/ ORGANISM: human
US-09-348-953-1

Query Match          50.9%; Score 258.5; DB 4; Length 82;
Best Local Similarity 54.3%; Pred. No. 2.7e-22;
Matches 44; Conservative 9; Mismatches 27; Indels 1; Gaps 1;

QY 3 DCMFGNGKGYRGKATVITGTPCQEWAAQEPHRSHTFIPGTNKWAGLEKNYCRNPDGDI 62
Db 2 ECKTGNGKNGYRGTMSTKNGITCQKWSSTSPHR-PFSPATHPSEGLEENYCRNPDNDPQ 60
QY 63 GPWCYTMNPKLFYDCDIPLC 83
Db 61 GPWCYTTDPKRYDYCDILEC 81

RESULT 10
US-08-612-788-8
/ Sequence 8, Application US/08612788
/ Patent No. 5837682
/ GENERAL INFORMATION:
/ APPLICANT: Folkman, M. Judah
/ APPLICANT: O'Reilly, Michael
/ APPLICANT: Cao, Yihai
/ APPLICANT: Sim, B. Kim Lee
/ TITLE OF INVENTION: Angiostatin Fragments and Method of Use
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta

/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Bovine
/ IMMEDIATE SOURCE:
/ CLONE: K1
/ SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-335-325-11

Query Match          50.9%; Score 258.5; DB 4; Length 79;
Best Local Similarity 57.5%; Pred. No. 2.6e-22;
Matches 46; Conservative 7; Mismatches 26; Indels 1; Gaps 1;

QY 4 CMFNGKGYRGKATVITGTPCQEWAAQEPHRSHTFIPGTNKWAGLEKNYCRNPDGDI 63
Db 1 CKTGNGKNGYRGTMSTKNGITCQKWSSTSPHR-PFSPATHPSEGLEENYCRNPDNDG 59
QY 64 PWCYTMNPKLFYDCDIPLC 83
Db 60 PWCITTDPRYDYCDILEC 79

RESULT 9
US-09-348-953-1
/ Sequence 1, Application US/09348953
/ Patent No. 6538103
/ GENERAL INFORMATION:
/ APPLICANT: Ji, Richard Wei-Dong
/ APPLICANT: Trill, Pamela A.
/ TITLE OF INVENTION: LYSINE BINDING FRAGMENTS OF ANGIOSTATIN
/ FILE REFERENCE: DB11sequences
/ CURRENT APPLICATION NUMBER: US/09/348,953
/ CURRENT FILING DATE: 1999-07-07
/ PRIOR APPLICATION NUMBER: 60/092,831
/ PRIOR FILING DATE: 1998-07-14
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 82
/ TYPE: PRT
/ ORGANISM: human
US-09-348-953-1

Query Match          50.5%; Score 256.5; DB 2; Length 79;
Best Local Similarity 55.0%; Pred. No. 4.4e-22;
Matches 44; Conservative 8; Mismatches 27; Indels 1; Gaps 1;

QY 4 CMFNGKGYRGKATVITGTPCQEWAAQEPHRSHTFIPGTNKWAGLEKNYCRNPDGDI 63
Db 1 CKTGNGKNGYRGTMSTKNGITCQKWSSTSPHR-PFSPATHPSEGLEENYCRNPDNDPQ 59
QY 64 PWCYTMNPKLFYDCDIPLC 83
Db 60 PWCITTDPRYDYCDILEC 79

RESULT 11
US-08-763-528A-2
/ Sequence 2, Application US/08763528A
/ Patent No. 5854221
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yihai
/ APPLICANT: Folkman, M. Judah
/ TITLE OF INVENTION: Endothelial Cell Proliferation Inhibitor
/ TITLE OF INVENTION: and Method of Use
/ NUMBER OF SEQUENCES: 6
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Jones & Askew, LLP
/ STREET: 191 Peachtree Street, 37th Floor
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: US
/ ZIP: 30303
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/763,528A
/ FILING DATE: 12-DEC-1996
```

```
/ CLASSIFICATION: 530
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05940-0251
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 79 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: internal
/ FEATURE:
/ NAME/KEY: Protein
/ LOCATION: 1..79
/ OTHER INFORMATION: /note= "Krangle 1 - Figure 3"
US-08-763-528A-2

Query Match 50.5%; Score 256.5; DB 2; Length 79;
Best Local Similarity 55.0%; Pred. No. 4.4e-22;
Matches 44; Conservative 8; Mismatches 27; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRCKKATVTGTPCOEWAAQEPHHSFTPIGTNKGWGLEKYNCRNPQGDING 63
DB 1 CKTGNGKNYRGMTSKTNGITCKQKWSSTSPHR-PRFSPATHPSGLEENYCRNPDPQ 59
QY 64 PWCYTMNPKLFYDCDIPLC 83
DB 60 PWCYTTDPKRYDYCDILEC 79

RESULT 12
US-09-066-028-8
; Sequence 8, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
```

```
/ TELEFAX: 404-818-3799
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 79 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ IMMEDIATE SOURCE:
/ CLONE: KI
/ US-09-066-028-8

Query Match 50.5%; Score 256.5; DB 3; Length 79;
Best Local Similarity 55.0%; Pred. No. 4.4e-22;
Matches 44; Conservative 8; Mismatches 27; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRCKKATVTGTPCOEWAAQEPHHSFTPIGTNKGWGLEKYNCRNPQGDING 63
DB 1 CKTGNGKNYRGMTSKTNGITCKQKWSSTSPHR-PRFSPATHPSGLEENYCRNPDPQ 59
QY 64 PWCYTMNPKLFYDCDIPLC 83
DB 60 PWCYTTDPKRYDYCDILEC 79

RESULT 13
US-09-335-325-8
; Sequence 8, Application US/09335325
; Patent No. 6521439
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
```

```

; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: K1
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-335-325-8

Query Match      50.5%; Score 256.5; DB 4; Length 79;
Best Local Similarity 55.0%; Pred. No. 4.4e-22;
Matches 44; Conservative 8; Mismatches 27; Indels 1; Gaps 1;

Qy 4 CMFGNGKGYRGKKATTVTGTCQEWAAQEPHHRSTFTPGTNKWAGLEKKNYCRNPDGDI 63
Db 1 CKTGNKNGYRGTSKNGITCKQWSSTSPHR-PFESPATHPSEGLEENYCRNPDNDPQG 59

Qy 64 PWCYTMNPKLFYDCDIPLC 83
Db 60 PWCYTTDPKRYDCDILEC 79

RESULT 14
US-08-612-788-10
; Sequence 10, Application US/08612788
; Patent No. 5837682
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Porcine
; IMMEDIATE SOURCE:
; CLONE: K1
US-08-612-788-10
```

```

Query Match      48.5%; Score 246.5; DB 2; Length 79;
Best Local Similarity 53.8%; Pred. No. 6e-21;
Matches 43; Conservative 7; Mismatches 29; Indels 1; Gaps 1;

Qy 4 CMFGNGKGYRGKKATTVTGTCQEWAAQEPHHRSTFTPGTNKWAGLEKKNYCRNPDGDI 63
Db 1 CKTGNKNGYRGTSKNGITCKQWSSTSPH-IPKYSPEKFLAGLEENYCRNPDNDKG 59

Qy 64 PWCYTMNPKLFYDCDIPIC 83
Db 60 PWCYTTDPETRFDYCDIPEC 79

RESULT 15
US-09-066-028-10
; Sequence 10, Application US/09066028
; Patent No. 6024688
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/612,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Porcine
; IMMEDIATE SOURCE:
; CLONE: K1
US-09-066-028-10

Query Match      48.5%; Score 246.5; DB 3; Length 79;
Best Local Similarity 53.8%; Pred. No. 6e-21;
Matches 43; Conservative 7; Mismatches 29; Indels 1; Gaps 1;

Qy 4 CMFGNGKGYRGKKATTVTGTCQEWAAQEPHHRSTFTPGTNKWAGLEKKNYCRNPDGDI 63
Db 1 CKTGNKNGYRGTSKNGITCKQWSSTSPH-IPKYSPEKFLAGLEENYCRNPDNDKG 59
```

Db 1 CKTGKKNYRGTTSTKXSGVICQKWSVSPH-IPXISPEKPEPLAGLEENYCRNPNDK 59

Qy 64 PWCYTNNRKLFDYCDIPLC 83

||||| : | ||||| |

Db 60 PWCYTTPETRFDYCDIPEC 79

Search completed: September 29, 2003, 08:50:31
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:49:20 : Search time 63 Seconds
(without alignments)
206.546 Million cell updates/sec

Title: US-10-088-548-8

Perfect score: 508

Sequence: 1 BQDCMFGNGYGRKATV.....YTNPKLFYCDIPLCASS 86

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 15130793 residues

Total number of hits satisfying chosen parameters: 198464

Minimum DB seq length: 0

Maximum DB seq length: 86

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

```
1: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pdb:
2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pdb:
3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pdb:
4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pdb:
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pdb:
6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pdb:
7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pdb:
8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pdb:
9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pdb:
10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pdb:
11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pdb:
12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pdb:
13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pdb:
14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pdb:
15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pdb:
16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pdb:
17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pdb:
18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pdb:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	412	81.1	80	9	US-09-753-064-6
2	412	81.1	80	12	US-10-267-137-8
3	405	79.7	79	9	US-09-753-064-1
4	387	76.2	80	9	US-09-761-120-47
5	274.5	54.0	79	9	US-09-761-120-9
6	274.5	54.0	79	10	US-09-335-325-9
7	274.5	54.0	79	15	US-10-131-241-9
8	258.5	50.9	79	9	US-09-761-120-11
9	258.5	50.9	79	10	US-09-335-325-11
10	258.5	50.9	79	15	US-10-131-241-11
11	256.5	50.5	79	9	US-09-753-064-2
12	256.5	50.5	79	9	US-09-761-120-8
13	256.5	50.5	79	10	US-09-335-325-8
14	256.5	50.5	79	12	US-10-267-137-7
15	256.5	50.5	79	15	US-10-131-241-8

```
16 246.5 48.5 79 9 US-09-761-120-10
17 246.5 48.5 79 10 US-09-335-325-10
18 246.5 48.5 79 15 US-10-131-241-10
19 241.5 47.5 79 10 US-09-335-325-7
20 241.5 47.5 79 15 US-10-131-241-7
21 240 47.2 78 9 US-09-761-120-22
22 240 47.2 78 10 US-09-335-325-22
23 240 47.2 78 15 US-10-131-241-22
24 232.5 45.8 79 9 US-09-761-120-7
25 231.5 45.6 79 12 US-10-267-137-10
26 231.5 45.6 80 12 US-10-267-137-1
27 222 43.7 78 9 US-09-753-064-5
28 222 43.7 78 9 US-09-761-120-17
29 222 43.7 78 9 US-09-761-120-23
30 222 43.7 78 10 US-09-335-325-17
31 222 43.7 78 10 US-09-335-325-23
32 222 43.7 78 12 US-10-267-137-9
33 222 43.7 78 15 US-10-131-241-17
34 222 43.7 78 15 US-10-131-241-23
35 220 43.3 78 9 US-09-761-120-16
36 220 43.3 78 9 US-09-761-120-20
37 220 43.3 78 10 US-09-335-325-16
38 220 43.3 78 10 US-09-335-325-20
39 220 43.3 78 15 US-10-131-241-16
40 220 43.3 78 15 US-10-131-241-20
41 217 42.7 78 9 US-09-761-120-12
42 217 42.7 78 10 US-09-335-325-12
43 217 42.7 78 15 US-10-131-241-12
44 216 42.5 78 9 US-09-761-120-14
45 216 42.5 78 10 US-09-335-325-13
```

ALIGNMENTS

RESULT 1

```
US-09-753-064-6
; Sequence 6, Application US/09753064
; Patent No. US20010016644A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; Folkmann, M. Judah
; TITLE OF INVENTION: Endothelial Cell Proliferation Inhibitor
; and Method of Use
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew, LLP
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/753,064
; FILING DATE: 29-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE: 12-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05940-0251
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
```

```
Sequence 10, Appl
Sequence 10, Appl
Sequence 10, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 22, Appl
Sequence 22, Appl
Sequence 22, Appl
Sequence 7, Appl
Sequence 10, Appl
Sequence 1, Appl
Sequence 5, Appl
Sequence 17, Appl
Sequence 23, Appl
Sequence 17, Appl
Sequence 23, Appl
Sequence 16, Appl
Sequence 20, Appl
Sequence 16, Appl
Sequence 20, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 14, Appl
Sequence 13, Appl
```

```
;
; LENGTH: 80 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..80
; OTHER INFORMATION: /note= "Kringle 5 - Figure 3"
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-753-064-6

Query Match      81.1%; Score 412; DB 9; Length 80;
Best Local Similarity 83.8%; Pred. No. 6.5e-39;
Matches 67; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 4 CMFGNGKGYRGKATVTGTCQEWAAQEPHRSHTFPGTNKAGLEKNCNRPDGDING 63
Db 1 CMFGNGKGYRGKATVTGTCQEWAAQEPHRSHTFPGTNKAGLEKNCNRPDGDVGG 60

QY 64 PWCYTMNPKLFYDCDIPLC 83
Db 61 PWCYTMNPKLYDCDVPQC 80

RESULT 2
US-10-267-137-8
; Sequence 8, Application US/10267137
; Publication No. US20030148950A1
; GENERAL INFORMATION:
; APPLICANT: Xin, Li
; APPLICANT: Li, Zai-Ping
; APPLICANT: Gan, Ren-bao
; APPLICANT: Zhou, Qing-wei
; TITLE OF INVENTION: KRINGLE DOMAIN 1 OF HUMAN HEPATOCYTE
; FILE REFERENCE: 52428200100
; CURRENT APPLICATION NUMBER: US/10/267,137
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 60/328,329
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 80
; TYPE: PRI
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence alignment of kringle domains of
; OTHER INFORMATION: plasminogen and HGF
US-10-267-137-8

Query Match      81.1%; Score 412; DB 12; Length 80;
Best Local Similarity 83.8%; Pred. No. 6.5e-39;
Matches 67; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 4 CMFGNGKGYRGKATVTGTCQEWAAQEPHRSHTFPGTNKAGLEKNCNRPDGDING 63
Db 1 CMFGNGKGYRGKATVTGTCQEWAAQEPHRSHTFPGTNKAGLEKNCNRPDGDVGG 60

QY 64 PWCYTMNPKLFYDCDIPLC 83
Db 61 PWCYTMNPKLYDCDVPQC 80

RESULT 3
US-09-753-064-1
; Sequence 1, Application US/09753064
; Patent No. US2001001664A1
```

```
;
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; TITLE OF INVENTION: Endothelial Cell Proliferation Inhibitor
; and Method of Use
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew, LLP
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/753,064
; FILING DATE: 29-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/763,528
; FILING DATE: 12-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05940-0251
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-753-064-1

Query Match      79.7%; Score 405; DB 9; Length 79;
Best Local Similarity 84.6%; Pred. No. 3.9e-38;
Matches 66; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 4 CMFGNGKGYRGKATVTGTCQEWAAQEPHRSHTFPGTNKAGLEKNCNRPDGDING 63
Db 1 CMFGNGKGYRGKATVTGTCQEWAAQEPHRSHTFPGTNKAGLEKNCNRPDGDVGG 60

QY 64 PWCYTMNPKLFYDCDIP 81
Db 61 PWCYTMNPKLYDCDVP 78

RESULT 4
US-09-761-120-47
; Sequence 47, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasmin
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
```

```
/ NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 47
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Kringle 5
US-09-761-120-47

Query Match          76.2%; Score 387; DB 9; Length 80;
Best Local Similarity 80.0%; Pred. No. 4.2e-36;
Matches 64; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 4 CMFGNGKGYRGKATVTGTPCOEWAAQEPHRSHTFIPGINKWAGLEKNCYCRNPDGNG 63
   ||||| ||||| || ||||| ||||| || || ||||| ||||| ||
Db 1 CMYGNKDYRGKTAAGTAAAGTFCQGWAAQEPHRSHTFTPTNPRADLEKNCYCRNPDGNG 60

QY 64 PWCYTMNPKLFEDYCDIPLC 83
   ||||| ||||| ||||| |||||
Db 61 PWCYTMNPKLFEDYCDIPLC 80

RESULT 5
US-09-761-120-9
; Sequence 9, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasminogen
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 9
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Macaca sp.
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Kringle 1
US-09-761-120-9

Query Match          54.0%; Score 274.5; DB 9; Length 79;
Best Local Similarity 58.8%; Pred. No. 1.7e-23;
Matches 47; Conservative 8; Mismatches 24; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATVTGTPCOEWAAQEPHRSHTFIPGINKWAGLEKNCYCRNPDGNG 63
   ||||| ||||| || ||||| ||||| || || ||||| ||||| ||
Db 1 CKTGNGKNYRGTMKTRGTGTCQKWSSTSPHR-PTFSPATHPSEGLENYCRNPDNGQG 59

QY 64 PWCYTMNPKLFEDYCDIPLC 83
   ||||| ||||| ||||| |||||
Db 60 PWCYTMNPKLFEDYCDIPLC 83

RESULT 6
US-10-131-241-9
; Sequence 9, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
```



```
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Rhesus monkey
US-10-131-241-9

Query Match          54.0%; Score 274.5; DB 15; Length 79;
Best Local Similarity 58.8%; Pred. No. 1.7e-23;
Matches 47; Conservative 8; Mismatches 24; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATVTGTCQEWAAQEPHRSHTFPGTNKWAGLEKNYCRNPDGDING 63
Db 1 CKTGNGKNTVGRGMSKTRIGTCQKWSSTSPH-PTFSPATHPSEGLEENYCRNPDNDGOG 59

QY 64 PWCYTMNPKLFYCDIPLC 83
Db 60 PWCYTTDPKRYDYCDIPEC 79

RESULT 8
US-09-761-120-11
; Sequence 11, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasminogen
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Bos sp.
; NAME/KEY: misc feature
; OTHER INFORMATION: Kringle 1
US-09-761-120-11

Query Match          50.9%; Score 258.5; DB 9; Length 79;
Best Local Similarity 57.5%; Pred. No. 1.1e-21;
Matches 46; Conservative 7; Mismatches 26; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATVTGTCQEWAAQEPHRSHTFPGTNKWAGLEKNYCRNPDGDING 63
Db 1 CKTGNGQTVRGTTAETKSGVTCQKWSATSPH-VPKFSPKFPFLAGLEENYCRNPDNDENG 59

QY 64 PWCYTMNPKLFYCDIPLC 83
Db 60 PWCYTTDPKRYDYCDIPEC 79

RESULT 9
US-09-335-325-11
; Sequence 11, Application US/09335325
; Patent No. US20020164717A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Bovine
; IMMEDIATE SOURCE:
; CLONE: K1
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-335-325-11

Query Match          50.9%; Score 258.5; DB 10; Length 79;
Best Local Similarity 57.5%; Pred. No. 1.1e-21;
Matches 46; Conservative 7; Mismatches 26; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATVTGTCQEWAAQEPHRSHTFPGTNKWAGLEKNYCRNPDGDING 63
Db 1 CKTGNGQTVRGTTAETKSGVTCQKWSATSPH-VPKFSPKFPFLAGLEENYCRNPDNDENG 59

QY 64 PWCYTMNPKLFYCDIPLC 83
Db 60 PWCYTTDPKRYDYCDIPEC 79

RESULT 10
US-10-131-241-11
; Sequence 11, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
```

```
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Bovine
US-10-131-241-11

Query Match          50.9%; Score 258.5; DB 15; Length 79;
Best Local Similarity 57.5%; Pred. No. 1.1e-21;
Matches 46; Conservative 7; Mismatches 26; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATTTGTPCOEWAAQEPHRHSTFIPGTNKWAGLEKNCVCRNPDDING 63
Db 1 CKTGNGKYRGTTATSGTCKNSATSPH-VPKFSPKPLAGLENYCRNPDDENG 59

QY 64 PWCYTMNPKLFYDCDIPLC 83
Db 60 PWCYTTDPKRYDCDIPEC 79

RESULT 11
US-09-753-064-2
; Sequence 2, Application US/09753064
; Patent No. US20010016644A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; TITLE OF INVENTION: Endothelial Cell Proliferation Inhibitor
; and Method of Use
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew, LLP
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/753,064
; FILING DATE: 29-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/763,528
; FILING DATE: 12-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05940-0251
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..79
; OTHER INFORMATION: /note= "Kringle 1 - Figure 3"
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
```

```
US-09-753-064-2

Query Match          50.5%; Score 256.5; DB 9; Length 79;
Best Local Similarity 55.0%; Pred. No. 1.8e-21;
Matches 44; Conservative 8; Mismatches 27; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATTTGTPCOEWAAQEPHRHSTFIPGTNKWAGLEKNCVCRNPDDING 63
Db 1 CKTGNGKNYRGTMKTKNGITCKWSSTSPHR-PREFSPATHPSSEGLENYCRNPDDPQG 59

QY 64 PWCYTMNPKLFYDCDIPLC 83
Db 60 PWCYTTDPKRYDCDILEC 79

RESULT 12
US-09-761-120-8
; Sequence 8, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasm
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Kringle 1
US-09-761-120-8

Query Match          50.5%; Score 256.5; DB 9; Length 79;
Best Local Similarity 55.0%; Pred. No. 1.8e-21;
Matches 44; Conservative 8; Mismatches 27; Indels 1; Gaps 1;

QY 4 CMFGNGKGYRGKATTTGTPCOEWAAQEPHRHSTFIPGTNKWAGLEKNCVCRNPDDING 63
Db 1 CKTGNGKNYRGTMKTKNGITCKWSSTSPHR-PREFSPATHPSSEGLENYCRNPDDPQG 59

QY 64 PWCYTMNPKLFYDCDIPLC 83
Db 60 PWCYTTDPKRYDCDILEC 79

RESULT 13
US-09-335-325-8
; Sequence 8, Application US/09335325
; Patent No. US20020164717A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Micheal
; APPLICANT: Cao, Yihai
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
```

OTHER INFORMATION: Sequence alignment of kringle domains of

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: September 29, 2003, 08:44:24 ; Search time 39 Seconds
(without alignments)
212.064 Million cell updates/sec

Title: US-10-088-548-8
Perfect score: 508
Sequence: 1 EQDCMFGNGKGYRGKATV.....YTMNPKLFDYCDIPLCASS 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168692 residues
Total number of hits satisfying chosen parameters: 29185

Minimum DB seq length: 0
Maximum DB seq length: 86

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76.*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54.5	10.7	64	1 NTSR5L	neurotoxin V - Egy
2	54.5	10.7	65	2 E82078	conserved hypothet
3	52.5	10.3	64	1 NTSR5M	neurotoxin V - sco
4	51	10.0	65	2 T03082	hypothetical prote
5	49.5	9.7	68	2 G82600	hypothetical prote
6	48.5	9.5	65	1 NTSREB	neurotoxin XI - sc
7	48.5	9.5	65	2 PQ0852	glycoprotein H - h
8	48	9.4	40	2 T35510	hypothetical prote
9	47.5	9.4	40	2 S21332	hypothetical prote
10	47.5	9.4	57	2 D81186	hypothetical prote
11	47	9.3	59	1 T6P5U	mambin - eastern J
12	47	9.3	77	2 AF2564	hypothetical prote
13	46.5	9.2	51	2 AF3350	hypothetical prote
14	46.5	9.2	62	2 A69034	hypothetical prote
15	46.5	9.2	73	2 F83824	hypothetical prote
16	46.5	9.2	86	2 T17313	hypothetical prote
17	46	9.1	71	2 A34768	ORF1 protein - Ori
18	45.5	9.0	65	2 PQ0850	glycoprotein H - h
19	45.5	9.0	65	2 PQ0848	hypothetical prote
20	45.5	9.0	65	2 AE2406	hypothetical prote
21	45	8.9	46	2 S67923	telomeric DNA bind
22	45	8.9	80	2 E97800	hypothetical prote
23	44.5	8.8	50	2 E82529	hypothetical prote
24	44.5	8.8	62	2 T02631	hypothetical prote
25	44	8.7	62	2 G64561	hypothetical prote
26	44	8.7	64	2 A35940	neurotoxin alpha-I
27	44	8.7	69	2 T14285	embryogenic callus
28	44	8.7	74	2 A53356	toxin CSRX-1 - wan
29	44	8.7	85	2 A39306	alpha insect toxin

30	44	8.7	86	2 F82496	hypothetical prote
31	43.5	8.6	49	2 S70093	hypothetical prote
32	43.5	8.6	63	2 AE2350	hypothetical prote
33	43.5	8.6	70	2 AC0718	probable cold shoc
34	43.5	8.6	86	2 G69452	molybdopterin conv
35	43	8.5	52	2 T32835	hypothetical prote
36	43	8.5	59	2 S39068	homeotic protein c
37	43	8.5	65	2 H72805	gpi1 protein - Myc
38	43	8.5	74	2 H84256	hypothetical prote
39	43	8.5	76	1 IHER2	high potential iro
40	42.5	8.4	58	2 F84303	50S ribosomal prot
41	42.5	8.4	66	2 B83515	hypothetical prote
42	42.5	8.4	81	2 A41441	intestinal trefoil
43	42.5	8.4	82	2 S37118	class II histocomp
44	42	8.3	59	2 AF2245	hypothetical prote
45	42	8.3	52	2 A12767	hypothetical prote

ALIGNMENTS

RESULT 1

NTSR5L
neurotoxin V - Egyptian scorpion
C;Species: Leiurus quinquestriatus (Egyptian scorpion)
C;Date: 30-Apr-1979 #sequence_revision 13-Jul-1981 #text_change 15-Oct-1996
C;Accession: A01741
R;Kopeyan, C.; Martinez, G.; Rochat, H.
FEBS Lett. 89, 54-58, 1978
A;Title: Amino acid sequence of neurotoxin V from the scorpion Leiurus quinquestriatus
A;Reference number: A01741; MUID:78191225; PMID:658402
A;Note: L. q. quinquestriatus
A;Accession: A01741
A;Molecule type: Protein
A;Residues: 1-64 <KOP>

C;Superfamily: scorpion neurotoxin
C;Keywords: blocked carboxyl end; neurotoxin; venom
F;12-63,16-36,22-46,26-48/Bisulfide bonds: #status Predicted
F;64/Modified site: blocked carboxyl end (Asn) (probably amidated) #status experiment

Query Match 10.7%; Score 54.5; DB 1; Length 64;
Best Local Similarity 34.0%; Pred. NO. 49;
Matches 16; Conservative 5; Mismatches 15; Indels 11; Gaps 3;

QY 35 RHSTFIPTN-----KWGLEKNYCR--NFDGDLNGWCYTMNPR 72
DB 10 KNTFFCGRMAYCNDECKKKGGSGFCOWASPYG--NACWYKLPDR 54

RESULT 2

E82078
conserved hypothetical protein VC2429 [imported] - Vibrio cholerae (strain N16961 se
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: E82078
R;Heidelberger, J.F.; Eissen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R
l. R.R.; Metakalos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: E82078
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-65 <HEI>

A;Cross-references: GB:AE004312; GB:AE003852; NID:g9656995; PIDN:AAF95572.1; GSPDB:GI
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC2429
A;Map position: 1

Query Match 10.7%; Score 54.5; DB 2; Length 65;
Best Local Similarity 31.7%; Pred. NO. 50;


```

A:Accession: PQ0851
A:Molecule type: DNA
A:Residues: 1-65 <GOM1>
A:Experimental source: isolate C3
A:Accession: PQ0854
A:Molecule type: DNA
A:Residues: 1-65 <GOM2>
A:Experimental source: isolate C9
A:Accession: PQ0853
A:Molecule type: DNA
A:Residues: 1-65 <GOM2>
A:Experimental source: strain C6
A:Superfamily: herpesvirus glycoprotein H
C:Keywords: glycoprotein; membrane protein

Query Match          9.5%; Score 48.5; DB 2; Length 65;
Best Local Similarity 20.0%; Pred. No. 2.2e+02;
Matches 14; Conservative 9; Mismatches 18; Indels 29; Gaps 4;

QY 23 TPCQEWAAQEPHRSHTFIPGKNWAGLEKYNCRNPDGDINGPWC-----YTMNPK 73
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 12 TPCYSW-----RPWTISNESECKN--GNSENPIVRPGITFNFTYKNDTR 54

QY 74 LFDYCDIPLC 83
      :: :| |
Db 55 IV--QVQPK 61

RESULT 8
T35510
Hypothetical protein SC6G10.01c - Streptomyces coelicolor (fragment)
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C:Accession: T35510
R:Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, March 1999
A:Reference number: Z21581
A:Accession: T35510
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-40 <SEE>
A:Cross-references: EMBL:AI049497; PIDN:CAB39855.1; GSPDB:GN000070; SCQEDB:SC6G10.01c
C:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCQEDB:SC6G10.01c

Query Match          9.4%; Score 48; DB 2; Length 40;
Best Local Similarity 47.8%; Pred. No. 1.5e+02;
Matches 11; Conservative 2; Mismatches 6; Indels 4; Gaps 1;

QY 21 TCTPCQEWAAQEPHRSHTFIPGT 43
      ||| | | | | | | | | | |
Db 7 SGTPIRGAANPPPHR-----PCT 25

RESULT 9
S21332
Hypothetical protein 2 - Agrobacterium tumefaciens
C:Species: Agrobacterium tumefaciens
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 30-Jun-2001
C:Accession: S21332
R:Paulus, F.; Canaday, J.; Otten, L.
submitted to the EMBL Data Library, October 1990
A:Description: Limited host range Ti plasmids; recent origin from wide host range Ti pla
A:Reference number: S21331
A:Accession: S21332
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-59 <PAU>
A:Cross-references: EMBL:X55075; NID:g39138; PIDN:CA38901.1; PID:g39140

Query Match          9.4%; Score 47.5; DB 2; Length 59;
Best Local Similarity 31.1%; Pred. No. 2.6e+02;

```

```

Matches 14; Conservative 4; Mismatches 16; Indels 11; Gaps 2;

QY 22 GTPCQEWAAQEPHRSHTFIPGKNWAGLEKYNCRNPDGDINGPW 65
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 16 GSPAQEPSSRRPHRS-----AGLLVYRCRPRSSSPIHAAW 50

RESULT 10
D81186
Hypothetical protein NM0542 [imported] - Neisseria meningitidis (strain MC58 serogr
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: D81186
R:Rettelin, H.; Saunders, N.J.; Heideberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ri, H.; Glin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masiugnani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: D81186
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-67 <TET>
A:Cross-references: GB:AE002410; GB:AE002098; NID:g7225766; PIDN:AAF40971.1; PID:g72;
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NM0542

Query Match          9.4%; Score 47.5; DB 2; Length 67;
Best Local Similarity 29.0%; Pred. No. 2.9e+02;
Matches 9; Conservative 5; Mismatches 14; Indels 3; Gaps 1;

QY 41 PGTNKNWAGLEKYN---CRNPDGDINGPWCTT 68
      || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 23 PGNQWNGHPRVFLPLCEGSGSVACPYCGT 53

RESULT 11
T6EPF5J
mambin - eastern Jameson's mamba
N:Alternate names: dendroaspin; short toxin S5C1
C:Species: Dendroaspis jamesoni kaimosae (eastern Jameson's mamba)
C:Date: 19-Feb-1984 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C:Accession: A42561; S50910; A01683
R:McDowell, R.S.; Dennis, M.S.; Louie, A.; Shuster, M.; Mulkerrih, M.G.; Lazarus, R.;
Biochemistry 31, 4766-4772, 1992
A:Title: Mambin, a potent glycoprotein IIB-IIIa antagonist and platelet aggregation :
A:Reference number: A42561; MUID:92273586; PMID:1591238
A:Accession: A42561
A:Molecule type: protein
A:Residues: 1-59 <WCD>
A:Cross-references: PIDN:AA22225.1; PID:g249694
A:Experimental source: venom
R:Jaseja, M.; Lu, X.; Williams, J.A.; Sutcliffe, M.J.; Kakkar, V.V.; Parslow, R.A.; I
Eur. J. Biochem. 226, 861-868, 1994
A:Title: (1)H-NMR assignments and secondary structure of dendroaspin, an RGD-contain:
A:Reference number: S50910; MUID:95112851; PMID:7813476
A:Accession: S50910
A:Molecule type: protein
R:Joubert, F.J.; Taljaard, N.
Biochim. Biophys. Acta 579, 228-233, 1979
A:Title: Some properties and the complete primary structures of two reduced and S-ca
A:Reference number: A90630; MUID:79232598; PMID:465532
A:Accession: A01683
A:Molecule type: protein
A:Residues: 1-15, 'CT', 18-27, 'RNI', 28-32, 34-59 <YOD>
C:Comment: the cell attachment motif may serve in inhibiting platelet aggregation.
C:Superfamily: snake toxin
C:Keywords: toxin; venom
F:43-45/Region: cell attachment (R-G-D) motif

```

F:3-22,17-37,39-51,52-57/Disulfide bonds: #status predicted

Query Match 9.3%; Score 47; DB 1; Length 59;
Best Local Similarity 30.5%; Pred. No. 2.9e+02;
Matches 18; Conservative 3; Mismatches 14; Indels 24; Gaps 4;

QY 15 KKATTVTGTPQOE-----WAAQEPHRSHTFPGTNKAGLEKNYCRNPDGNGINGPWC 66
 ||||| || || || ||
Db 10 KPPTTET---COEDSCYNIW-----TFDNIIRGCG-----CFTPRGDMFGPIC 51
 ||||| || || || ||

RESULT 12
AF2564
hypothetical protein asr8513 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC71206
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AF2564
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriquchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF2564
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-77 <KUR>
A:Cross-references: GB:AP003604; PIDN:BAB77432.1; PID:g17134876; GSPDB:GN00183
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: asr8513
A:Genome: plasmid

Query Match 9.3%; Score 47; DB 2; Length 77;
Best Local Similarity 33.3%; Pred. No. 3.8e+02;
Matches 9; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

QY 40 IPGTNKWAGLEKNYCRNPDGNGINGPWC 66
 || || || ||
Db 40 VKGKCHPSPVGFCCPNDGRTSSGFC 66

RESULT 13
AF3350
hypothetical protein BMEI0788 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AF3350
R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AF3350
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-51 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL51969.1; PID:g17982729; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI0788
A:Map position: I

Query Match 9.2%; Score 46.5; DB 2; Length 51;
Best Local Similarity 31.4%; Pred. No. 2.9e+02;
Matches 11; Conservative 4; Mismatches 13; Indels 7; Gaps 1;

QY 24 PCQEWAAQEPHRSHTFPGTNKAGLEKNYCRNPD 58
 |||: :|||:| || ||
Db 5 PCTDAGHERPHRLSS-----AHKATNLCRRPE 32

RESULT 14

A69034
hypothetical protein MTH125 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: A69034
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T
.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,
K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: A69034
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-62 <MTH>
A:Cross-references: GB:AE000802; GB:AE000666; NID:g2621163; PIDN:AAB84631.1; PID:g262
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH125
A:Start codon: TTG

Query Match 9.28; Score 46.5; DB 2; Length 62;
Best Local Similarity 42.9%; Pred. No. 3.5e+02;
Matches 12; Conservative 2; Mismatches 13; Indels 1; Gaps 1;

QY 51 KNYCRNPDGNGINGPWC-YTMNPKLFY 77
 |||:| || ||
Db 35 KYCRHPRGPGNPSHCYFRDRKVDGY 62

RESULT 15
F83824
hypothetical protein BH1398 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: F83824
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: F83824
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-73 <STPO>
A:Cross-references: GB:AF001511; GB:BA000004; NID:g10173727; PIDN:BA805117.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1398

Query Match 9.28; Score 46.5; DB 2; Length 73;
Best Local Similarity 22.2%; Pred. No. 4.1e+02;
Matches 14; Conservative 4; Mismatches 28; Indels 17; Gaps 2;

QY 25 CQEWAAQ-----EPHRSHTFPGTNKAGLEKNYCRNPDGNGINGPWCYTMN 70
 |||: :|||:| || ||
Db 13 CMYWGGENEHCADALIVEIDANQHRYDMEAGREPYATAEKREYAKSSGE---TCCYIFR 69
 |||: :|||:| || ||
QY 71 PRK 73
 |||: :|||:| ||
Db 70 PKK 72

Search completed: September 29, 2003, 08:49:58
JOB time : 42 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	54.5	10.7	64	1	SCX5_LEIQH	P01481 leirurus qui
2	54.5	10.7	65	1	Y029_VIBCH	Q9xpel vibrio chol
3	54.5	10.7	82	1	CX06_CONTE	Q9xzl3 conus texti
4	52.5	10.3	64	1	SCX5_ANDMA	P01482 androctonus
5	49	9.6	67	1	GGBD_HUMAN	Q9p2w3 homo sapien
6	49	9.6	67	1	GGBD_MOUSE	Q9jmf3 mus musculus
7	48.5	9.5	64	1	YF29_VIBPA	Q871t2 vibrio para
8	48.5	9.5	65	1	SCXB_BUTOC	P01486 buthus occi
9	47	9.3	59	1	MAMB_DSNJA	P20375 dendroaspis
10	47	9.3	65	1	SCBK_MESMA	P82615 mesobuthus
11	46.5	9.2	85	1	SC15_MESMA	Q9ngn8 mesobuthus
12	46.5	9.2	85	1	SCAL_MESMA	Q9gyx2 mesobuthus
13	46	9.1	41	1	S482_HYSGI	P56854 hystero크라
14	46	9.1	61	1	TX51_DSNJA	P01413 dendroaspis
15	46	9.1	64	1	YG19_VIBVU	Q86c30 vibrio vuln
16	44	8.7	64	1	SCX3_LEIQH	P01487 leirurus qui
17	44	8.7	65	1	SIX1_BUTOC	P55902 buthus occi
18	44	8.7	74	1	TXCL_CUPSA	F81694 cupennius
19	44	8.7	85	1	SCX4_LEIQH	P17728 leirurus qui
20	43.5	8.6	70	1	CSPH_SALTY	Q33793 salmonella
21	43	8.5	65	1	VG51_BPMD2	O64241 mycobacteri
22	43	8.5	68	1	NLT2_PRIAR	P82353 prunus armel
23	43	8.5	76	1	HPI2_ECTHA	P04169 ectothiorbo
24	42.5	8.4	58	1	RL37_HALNI	Q9nps3 halobacteri
25	42.5	8.4	76	1	IBX0_CONTE	Q9xzk8 conus texti
26	42	8.3	64	1	IBB1_COILA	P07679 coix lachry
27	42	8.3	76	1	SPY1_CHICK	Q9pt11 gallus gall
28	42	8.3	84	1	SCXB_MESMA	Q9n682 mesobuthus
29	42	8.3	84	1	Y463_CAUCR	Q9aax9 caulobacter
30	42	8.3	85	1	Y4XH_BACSU	Q04811 bacillus su
31	41.5	8.2	49	1	R332_LISMO	Q94f31 listeria mon
32	41.5	8.2	84	1	TFF1_HUMAN	P04155 homo sapien
33	41	8.1	64	1	SCX4_MESMA	P59328 mesobuthus

DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DE Hypothetical zinc-binding UPF0243 protein VC2429.
 GN VC2429.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=656;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 EX MEDLINE=20406833; PubMed=10953301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
 RT *cholerae*.";
 RL Nature 406:477-483(2000).
 CC -!- COPACATOR: Binds 1 zinc ion (By similarity).
 CC -!- SIMILARITY: Belongs to the UPF0243 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE004312; AAF95572.1; -;
 DR PIR; E82078; E82078.
 DR TIGR; VC2429; -;
 DR HAMAP; MF_00649; -; 1.
 DR InterPro; IPR005584; DUF329.
 DR Pfam; PF03884; DUF329; 1.
 KW Hypothetical protein; Zinc; Metal-binding; Complete proteome.
 FT METAL 10 10 ZINC (BY SIMILARITY).
 FT METAL 13 13 ZINC (BY SIMILARITY).
 FT METAL 29 29 ZINC (BY SIMILARITY).
 FT METAL 33 33 ZINC (BY SIMILARITY).
 FT METAL 33 33 ZINC (BY SIMILARITY).
 SQ SEQUENCE 65 AA; 7397 MW; 2ED887C1D1D8AD90 CRC64;
 Query Match 10.7%; Score 54.5; DB 1; Length 65;
 Best Local Similarity 31.7%; Pred. No. 15;
 Matches 20; Conservative 2; Mismatches 24; Indels 17; Gaps 3;
 QY 15 KKATTVGTGTPC---QEWAAQEPHRSHTFPGTNK-----WAGLEKNCYCRNPDGGIN 62
 DB 3 KKLITVKPCRGDGVWVGQSPHR-----PCKSQQCMIDFGWAEDEKAIPGAPDMSDS 57
 QY 63 GFW 65
 DB 58 DGW 60
 RESULT 3
 CX06_CONTE STANDARD; PRT; 82 AA.
 AC QXZL3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Hypothetical zinc-binding UPF0243 protein VC2429.
 GN VC2429.
 OS Omega-type conotoxin Tx06 precursor.
 DE Tx06.
 GN Conus textile (Cloth-of-gold cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
 OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.

OX NCBI_TaxID=6494;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Venom duct;
 RA MEDLINE=20037955; PubMed=10573284;
 RX Lu B.-S., Yu F., Zhao D., Huang P.-T., Huang C.-F.;
 RT "Conopeptides from *Conus striatus* and *Conus textile* by cDNA
 RT cloning.";
 RL Peptides 20:1139-1144(1999).
 CC -!- FUNCTION: Omega-conotoxins act at presynaptic membranes, they bind
 CC and block voltage-sensitive calcium channels (VSCC) (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
 CC -!- SIMILARITY: BELONGS TO THE O-SUPERFAMILY OF CONOTOXINS. OMEGA-TYPE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF146359; AAD31919.1; -;
 DR InterPro; IPR004214; Conotoxin.
 DR Pfam; PF02950; Conotoxin; 1.
 KW Toxin; Neurotoxin; Presynaptic neurotoxin; Ionic channel inhibitor;
 KW Calcium channel inhibitor; Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT PROPEP 23 50 OMEGA-TYPE CONOTOXIN TX06.
 FT PEPTIDE 51 82 BY SIMILARITY.
 FT DISULFID 53 71 BY SIMILARITY.
 FT DISULFID 60 76 BY SIMILARITY.
 FT DISULFID 70 81 BY SIMILARITY.
 SQ SEQUENCE 82 AA; 9305 MW; 1739D2531B865860 CRC64;
 Query Match 10.7%; Score 54.5; DB 1; Length 82;
 Best Local Similarity 30.8%; Pred. No. 19;
 Matches 12; Conservative 4; Mismatches 6; Indels 17; Gaps 1;
 QY 47 AGLEKNCYCRNPDGNGWCTMNPRLDYCDIPLCAS 85
 DB 46 AKLEKNCY-----QEKWDYCPVPLGSG 67
 RESULT 4
 SCX5_ANDMA STANDARD; PRT; 64 AA.
 AC P01482;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neurotoxin V (AaH V) (AaH5).
 OS Androctonus mauretanicus mauretanicus (Scorpion).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
 OC Butioidae; Metidae; Androctonus.
 OX NCBI_TaxID=6860;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 EX MEDLINE=85193276; PubMed=3992595;
 RA ROSSO J.P., KOCHAT H.;
 RT "Characterization of ten proteins from the venom of the Moroccan
 RT scorpion *Androctonus mauretanicus mauretanicus*, six of which are
 RT toxic to the mouse.";
 RL Toxicon 23:113-125(1985).
 CC -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
 CC of the activated channels, thereby blocking neuronal transmission.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.

```

CC      ALPHA-TOXIN SUBFAMILY.
DR      PIR: A01742; NTRSM.
DR      HSSP: F01484; 1AHO.
DR      InterPro: IPR003614; Knott1.
DR      InterPro: IPR001219; Neurotoxin.
DR      InterPro: IPR002061; Scorpion_toxinL.
DR      Pfam: PF00537; toxin_3; 1.
DR      PRINTS: PR00284; TOXIN.
DR      ProDom: PD00908; scorpion_toxinL; 1.
DR      SMART: SM00505; Knott1; 1.
KW      Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW      Amidation.
FT      DISULFID 12 63 BY SIMILARITY.
FT      DISULFID 16 36 BY SIMILARITY.
FT      DISULFID 22 46 BY SIMILARITY.
FT      DISULFID 26 48 BY SIMILARITY.
FT      MOD_RES 64 64 AMIDATION.
SQ      SEQUENCE 64 AA; 7301 MW; 3D03A733534CD866 CRC64;

Query Match 10.3%; Score 52.5; DB 1; Length 64;
Best Local Similarity 36.4%; Pred.No. 25;
Matches 16; Conservative 3; Mismatches 14; Indels 11; Gaps 3;

QY      38 TFIPGTRN-----KWAGLEKNYCR--NPDGNGINGFWCYTMNPR 72
        || || || || || || || || || || || || || || || ||
        13 TFGGRNAYCDDECKKGSGSYGQWASPYG--NACWCYKLPDR 54

Db

RESULT 5
GBGD_HUMAN
ID      GBGD_MOUSE STANDARD; PRT; 67 AA.
AC      Q9PJW3; Q9UJ3;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Guanine nucleotide-binding protein G(i)/G(s)/G(o) gamma-13 subunit.
GN      GNG13.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=20040713; PubMed=10570481;
RA      Huang L., Shanker Y.G., Dubauskaite J., Zheng J.Z., Yan W.,
RA      Rosenzweig S., Spielman A.I., Max M., Margolske R.F.;
RT      "Ggamma13 colocalizes with gustducin in taste receptor cells and
RT      mediates IP3 responses to bitter denatonium.";
RL      Nat. Neurosci. 2:1055-1062(1999).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      TISSUE=Brain;
RA      Inoue S., Sano H., Ohta M.;
RT      "E.coli toxicity assay: a novel expression screening method for
RT      isolation of mammalian genes with membrane-associated domains or ATP
RT      binding/ATPase domains.";
RL      Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=2109610; PubMed=11157797;
RA      Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,
RA      Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,
RA      Higgs D.R.;
RT      "sequence, structure and pathology of the fully annotated terminal 2
RT      MB of the short arm of human chromosome 16.";
RL      Hum. Mol. Genet. 10:339-352(2001).
RN      [4]
RP      SEQUENCE FROM N.A.
RX      PubH H.L. III, Ikeda S.R., Aronstam R.S.;
RA      Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN      [5]
RP      SEQUENCE FROM N.A.
RA      Frankland J.;

```

```

RL      Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC      -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC      INVOLVED AS A MODULATOR OR TRANSDUCER IN VARIOUS TRANSMEMBRANE
CC      SIGNALING SYSTEMS. THE BETA AND GAMMA CHAINS ARE REQUIRED FOR THE
CC      GTPASE ACTIVITY, FOR REPLACEMENT OF GDP BY GTP, AND FOR G PROTEIN-
CC      EFFECTOR INTERACTION.
CC      -!- SUBUNIT: G proteins are composed of 3 units (alpha, beta and
CC      gamma).
CC      -!- SIMILARITY: BELONGS TO THE G PROTEIN GAMMA FAMILY.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      EMBL: AY029486; RAA40269.1; -.
DR      EMBL: AB030207; BAA92768.1; -.
DR      EMBL: AE006465; RAK61257.1; -.
DR      EMBL: AF493880; RAA12594.1; -.
DR      EMBL: AL031033; CAB33059.1; ALT_INIT.
DR      Genew; HGNC:14131; GNG13.
DR      MIM: 607298; -.
DR      InterPro: IPR001770; G-gamma.
DR      Pfam: PF00631; G-gamma; 1.
DR      PRINTS: PR00321; GPROTEING.
DR      ProDom: PD003783; G-gamma; 1.
DR      SMART: SM00224; GGL; 1.
DR      PROSITE: PS50058; G-PROTEIN_GAMMA; 1.
KW      Transducer; Prenylation; Lipoprotein; Multigene family.
FT      LIPID 64 64 FARNESYL (BY SIMILARITY).
FT      PROPEP 65 67 REMOVED IN MATURE FORM (BY SIMILARITY).
SQ      SEQUENCE 67 AA; 7949 MW; 080E0C348F70327 CRC64;

Query Match 9.6%; Score 49; DB 1; Length 67;
Best Local Similarity 41.9%; Pred.No. 64;
Matches 13; Conservative 2; Mismatches 14; Indels 2; Gaps 1;

QY      37 STPIPTKNKA--GLEKNYCNPDGNGINGPW 65
        | || || || || || || || || || ||
        28 SKTIPELLKWTEDSGIPKDFLNPULMKNPW 58

Db

RESULT 6
GBGD_MOUSE
ID      GBGD_MOUSE STANDARD; PRT; 67 AA.
AC      Q9JMF3;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Guanine nucleotide-binding protein G(i)/G(s)/G(o) gamma-13 subunit.
GN      GNG13.
OS      Mus musculus (Mouse)
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=C57BL/6J;
RA      Huang L., Shanker Y.G., Dubauskaite J., Zheng J.Z., Yan W.,
RA      Rosenzweig S., Spielman A.I., Max M., Margolske R.F.;
RT      "Ggamma13 colocalizes with gustducin in taste receptor cells and
RT      mediates IP3 responses to bitter denatonium.";
RL      Nat. Neurosci. 2:1055-1062(1999).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      TISSUE=Brain;
RA      Inoue S., Sano H., Ohta M.;
RT      "Growth suppression of Escherichia coli by induction of expression of

```

ID	YP29_VTBPA	STANDARD;	PRT;	64 AA.
AC	Q87LT2;			
DT	15-SEP-2003 (Rel. 42, Created)			
DD	15-SEP-2003 (Rel. 42, Last sequence update)			
DE	15-SEP-2003 (Rel. 42, Last annotation update)			
DN	Hypothetical zinc-binding UPF0243 protein VP2529.			
GS	VP2529.			
OS	Vibrio parahaemolyticus.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;			
OC	Vibrionaceae; Vibrio.			
OX	NCBI_TaxID=670;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=RIMD 2210633 / Serotype O3:K6;			
RC	MEDLINE=2508454; PubMed=12620739;			
RA	Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,			
RA	Iijima Y., Najiima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,			
RA	Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;			
RT	Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism			
RT	distinct from that of V. cholerae.;			
RT	Lancet 361:743-749(2003).			
CC	-I- COFACTOR: Binds 1 zinc ion (By similarity).			
CC	-I- SIMILARITY: Belongs to the UPF0243 family.			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC				
CC	EMBL; AF005081; BAG60792.1; -			
DR	HMAP; MF 00649; -; 1.			
KW	Hypothetical protein; Zinc; Metal-binding; Complete proteome.			
FT	METAL 9 9 ZINC (BY SIMILARITY).			
FT	METAL 12 12 ZINC (BY SIMILARITY).			
FT	METAL 28 28 ZINC (BY SIMILARITY).			
FT	METAL 32 32 ZINC (BY SIMILARITY).			
SQ	SEQUENCE 64 AA: 7187 MW; AD7939573E9F9F9F9 CRG64;			
Query Match	9.5%;	Score 48.5;	DB 1;	Length 64;
Best Local Similarity	29.0%;	Pred. No. 69;		
Matches	18;	Conservative	2;	Mismatches 25; Indels 17; Gaps 3
QY	16 KATVITGTPC---QEWAQEPHERSTFIPTNK-----WAGLEKNYCRNPDGSDING 63			
DB	3 KITIVCPQCGTDVWEGQSPR-----PFCCKKKCMIDFGWADEENAIAGAPDMSD 57			
QY	64 PW 65			
DB	58 GW 59			
Query Match	9.5%;	Score 48.5;	DB 1;	Length 64;
Best Local Similarity	29.0%;	Pred. No. 69;		
Matches	18;	Conservative	2;	Mismatches 25; Indels 17; Gaps 3
QY	16 KATVITGTPC---QEWAQEPHERSTFIPTNK-----WAGLEKNYCRNPDGSDING 63			
DB	3 KITIVCPQCGTDVWEGQSPR-----PFCCKKKCMIDFGWADEENAIAGAPDMSD 57			
QY	64 PW 65			
DB	58 GW 59			
Query Match	9.5%;	Score 48.5;	DB 1;	Length 64;
Best Local Similarity	29.0%;	Pred. No. 69;		
Matches	18;	Conservative	2;	Mismatches 25; Indels 17; Gaps 3
QY	16 KATVITGTPC---QEWAQEPHERSTFIPTNK-----WAGLEKNYCRNPDGSDING 63			
DB	3 KITIVCPQCGTDVWEGQSPR-----PFCCKKKCMIDFGWADEENAIAGAPDMSD 57			
QY	64 PW 65			
DB	58 GW 59			
Query Match	9.5%;	Score 48.5;	DB 1;	Length 64;
Best Local Similarity	29.0%;	Pred. No. 69;		
Matches	18;	Conservative	2;	Mismatches 25; Indels 17; Gaps 3
QY	16 KATVITGTPC---QEWAQEPHERSTFIPTNK-----WAGLEKNYCRNPDGSDING 63			
DB	3 KITIVCPQCGTDVWEGQSPR-----PFCCKKKCMIDFGWADEENAIAGAPDMSD 57			
QY	64 PW 65			
DB	58 GW 59			
Query Match	9.5%;	Score 48.5;	DB 1;	Length 64;
Best Local Similarity	29.0%;	Pred. No. 69;		
Matches	18;	Conservative	2;	Mismatches 25; Indels 17; Gaps 3
QY	16 KATVITGTPC---QEWAQEPHERSTFIPTNK-----WAGLEKNYCRNPDGSDING 63			
DB	3 KITIVCPQCGTDVWEGQSPR-----PFCCKKKCMIDFGWADEENAIAGAPDMSD 57			
QY	64 PW 65			
DB	58 GW 59			
Query Match	9.5%;	Score 48.5;	DB 1;	Length 64;
Best Local Similarity	29.0%;	Pred. No. 69;		
Matches	18;	Conservative	2;	Mismatches 25; Indels 17; Gaps 3
QY	16 KATVITGTPC---QEWAQEPHERSTFIPTNK-----WAGLEKNYCRNPDGSDING 63			
DB	3 KITIVCPQCGTDVWEGQSPR-----PFCCKKKCMIDFGWADEENAIAGAPDMSD 57			
QY	64 PW 65			
DB	58 GW 59			
Query Match	9.5%;	Score 48.5;	DB 1;	Length 64;
Best Local Similarity	29.0%;	Pred. No. 69;		
Matches	18;	Conservative	2;	Mismatches 25; Indels 17; Gaps 3
QY	16 KATVITGTPC---QEWAQEPHERSTFIPTNK-----WAGLEKNYCRNPDGSDING 63			
DB	3 KITIVCPQCGTDVWEGQSPR-----			

DE	Buthus occitanus tunetanus (Common European scorpion).
OS	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC	Euthoidae; Buthidae; Buthus.
OX	NCBI_TaxID=6871;
RN	[1]
RP	SEQUENCE.
RC	TISSUE=Venom;
RX	MEDLINE=84224814; PubMed=6729843;
RA	Martin M.-P., Rochat H.;
RT	"Purification of thirteen toxins active on mice from the venom of the
RT	North African scorpion Buthus occitanus tunetanus.";

```

CC THE BINDING OF PURIFIED PLATELET FIBRINOGEN RECEPTOR GPIIb-IIIa
CC TO IMMOBILIZED FIBRINOGEN.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: Belongs to the snake toxin family.
DR PIR: A42561, T6PESJ.
DR PDB: 1DRS; 20-DEC-94.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam: PF00087; toxin; 1.
DR ProDom: PD000206; Snake_toxin; 1.
DR PROSITE: PS00272; SNAKE_TOXIN; FALSE NEG.
KW Blood coagulation; Cell adhesion; Toxin; 3D-structure.
FT DISULFID 3 22
FT DISULFID 17 37
FT DISULFID 39 51
FT DISULFID 52 57
FT SITE 43 45
FT STRAND 2 4
FT TURN 8 9
FT STRAND 14 16
FT STRAND 22 24
FT TURN 28 31
FT STRAND 36 37
FT STRAND 50 52
FT SEQUENCE 59 AA; 6754 MW; 20DC6A59DF9E41 CRC64;

Query Match 9.3%; Score 47; DB 1; Length 59;
Best Local Similarity 30.5%; Pred. No. 93;
Matches 18; Conservative 3; Mismatches 14; Indels 24; Gaps 4;

QY 15 KKATVTGTPCGE-----WAAQEPHRHSTFTPGTNKWAGLEKNYCRNPDGSDINGPWC 66
Db 10 KPPTET---CQEDSCYKNW-----TETNIRRGOG-----CFTPRGDMPGPYC 51

RESULT 10
ID SCBK_MESMA STANDARD; PRT; 65 AA.
AC P82815;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eukatoxin.
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Mesobuthus.
OX NCBI_TaxID=34649;
RN [1]
RC TISSUE=venom;
RX MEDLINE=21211325; PubMed=11311230;
RA Srinivasan K.N., Nirthanan S., Sasaki T., Sato K., Cheng B.,
RA Gwee M.C.B., Kini R.M., Gopalakrishnakone P.;
RA "Functional site of Eukatoxin, an alpha-type sodium channel
RT neurotoxin from the Chinese scorpion (Buthus martensii Karsch) venom:
RT probable role of the (52)PKVP(56) loop.";
RL FEBS Lett. 494:145-149(2001).
CC -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC of the activated channels, thereby blocking neuronal transmission.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC ALPHA-TOXIN SUBFAMILY.
DR HSP; P17728; ILQ1.
DR InterPro: IPR003614; Knot1.
DR InterPro: IPR001219; Neurotoxin.
DR InterPro: IPR002061; Scorpion_toxinL.
DR Pfam: PF00537; toxin_3; 1.
DR PRINTS: PR00284; TOXIN.
DR ProDom: PD000908; Scorpion_toxinL; 1.
DR SMART: SM00505; knot1; 1.
KW Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor; Toxin.
FT DISULFID 12
FT SEQUENCE 63 BY SIMILARITY

```

```

FT DISULFID 16 36 BY SIMILARITY.
FT DISULFID 22 46 BY SIMILARITY.
FT DISULFID 26 48 BY SIMILARITY.
SQ SEQUENCE 65 AA; 7226 MW; F2AYDV60E7B959DA CRC64;

Query Match
Best Local Similarity 9.3%; Score 47; DB 1; Length 65;
Matches 12; Conservative 7; Mismatches 20; Indels 8; Gaps 2;

QY 35 RHSTFTPGINKWA-----GLKNYCRNPDGDPGNGWYTMNPKL 74
DQ 10 KNCAYFCGRNAYCDECIINGAESGYQQAGVYGNACWCYKL-PDKV 55

RESULT 11
SC15_MESMA
ID SC15_MESMA STANDARD; PRT; 85 AA.
AC Q9GNG8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Alpha-neurotoxin TX15 precursor.
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Mesobuthus.
OX NCBI_TaxID=34649;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=venom gland.
RX MEDLINE=20317244; PubMed=10858508;
RA Zhu S.-Y., Li W.-X., Zeng X.-C., Liu H., Jiang D.-H., Mao X.;
RT "Nine novel precursors of Buthus martensii scorpion alpha-toxin
homologues."
RL Toxicon 38:1653-1661(2000).
CC -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
of the activated channels, thereby blocking neuronal transmission.
CC This toxin is active against mammals (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC ALPHA-TOXIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
EMBL; AF163017; AAC39643.1; -.
DR EMBL; AF163016; AAC39642.1; -.
DR HSP; P17728; ILQI.
DR InterPro; IPR003614; Knot1.
DR InterPro; IPR001219; Neurotoxin.
DR InterPro; IPR002061; Scorpion_toxinL.
DR Pfam; PF00537; toxin_3; 1.
DR PRINTS; PR00284; TOXIN.
DR ProDom; PD000908; Scorpion_toxinL; 1.
DR SMART; SM00505; Knot1; 1.
DR Toin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW Signal.
FT SIGNAL. 1 19 POTENTIAL.
FT CHAIN 20 85 ALPHA-NEUROTOXIN TX15.
FT DISULFID 31 82 BY SIMILARITY.
FT DISULFID 35 55 BY SIMILARITY.
FT DISULFID 41 65 BY SIMILARITY.
FT DISULFID 45 67 BY SIMILARITY.
SQ SEQUENCE 85 AA; 9398 MW; 3312B422E1C57E1 CRC64;

Query Match
Best Local Similarity 9.2%; Score 46.5; DB 1; Length 85;
Matches 11; Conservative 4; Mismatches 7; Indels 11; Gaps 2;

FT SIGNAL. 1 19 POTENTIAL.
FT CHAIN 20 85 ALPHA-NEUROTOXIN TX15.
FT DISULFID 31 82 BY SIMILARITY.
FT DISULFID 35 55 BY SIMILARITY.
FT DISULFID 41 65 BY SIMILARITY.
FT DISULFID 45 67 BY SIMILARITY.
SQ SEQUENCE 85 AA; 9398 MW; 3312B422E1C57E1 CRC64;

Query Match
Best Local Similarity 9.2%; Score 46.5; DB 1; Length 85;
Matches 11; Conservative 4; Mismatches 7; Indels 11; Gaps 2;

QY 42 GTNKWAGLEKNCYCRNPDGDPGNGWYTMNPKL 74
DQ 53 GYCQWAGVYGNAC-----WCYKL-PDKV 74

RESULT 13
S482_HYSGI
ID S482_HYSGI STANDARD; PRT; 41 AA.
AC P56854;

```

```

DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Toxin SNX-482.
OS Hysterochrates gigas (African tarantula).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Mygalomorphae; Theraphosidae; Hysterochrates.
OX NCBI_TaxID=118972;
RN [1]
RN SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=99017889; PubMed=9799496;
RA Newcomb R., Szoke B., Palma A., Wang G., Chen X.H., Hopkins W.,
RA Cong R., Miller J., Urge L., Parczyk-Hornoch K., Loo J.A., Dooley D.J.,
RA Radasdi L., Tsien R.W., Lemos J., Miljanich G.;
RT "selective peptide antagonist of the class E calcium channel from the
RT venom of the tarantula Hysterochrates gigas.";
RL Biochemistry 37:15353-15362(1998).
CC -!- FUNCTION: POTENT AND SELECTIVE BLOCKER OF THE CLASS E CALCIUM
CC CHANNEL.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: BELONGS TO THE SPIDER POTASSIUM CHANNEL INHIBITORY
CC TOXIN FAMILY.
DR HSP; P56852; 1D1H.
KW Toxin; Neurotoxin; Ionic channel inhibitor; Calcium channel inhibitor.
SQ SEQUENCE 41 AA; 4501 MW; C5D3CC96D3A0CF74 CRC64;

Query Match 9.1%; Score 46; DB 1; Length 41;
Best Local Similarity 33.3%; Pred. No. 83;
Matches 11; Conservative 4; Mismatches 16; Indels 2; Gaps 1;

QY 48 GLEKXNCRNPDG-DINGPCWYTMNPKLPDYC 78
Dl : : : | | | | | : | : | : | | | |
Db 1 GVDKAGCYMFGGSVNDCCPRLGCHSLFSYC 33

RESULT 14
TX51_DENJA
ID TX51_DENJA STANDARD; PRT; 61 AA.
AC P01413;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Toxin S5C1.
OS Dendroaspis jamesoni kaimosae (Eastern Jameson's mamba).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosaurs; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Elapidae; Elapinae; Dendroaspis.
OX NCBI_TaxID=8619;
RN [1]
RN SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=79232598; PubMed=465532;
RA Joubert F.O., Taljaard N.;
RT "Some properties and the complete primary structures of two reduced
RT and S-carboxymethylated polypeptides (S5C1 and S5C10) from
RT Dendroaspis jamesoni kaimosae (Jameson's mamba) venom.";
RL Biochim. Biophys. Acta 579:228-233(1979).
CC -!- FUNCTION: INHIBITS ADP-INDUCED PLATELET AGGREGATION AND INHIBITS
CC THE BINDING OF PURIFIED PLATELET FIBRINOGEN RECEPTOR GPIIb-IIIa
CC TO IMMOBILIZED FIBRINOGEN (BY SIMILARITY TO NAMBIN).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: Belongs to the snake toxin family.
DR HSP; P28375; 1DPS.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam: PF00067; toxin; 1.
DR PRODOM: PD000206; Snake_toxin; 1.
DR PROSITE: PS00272; SNAKE_TOXIN; FALSE_NEG.
KW Blood coagulation; Cell adhesion; Toxin.
FT DISULFID 3 22 BY SIMILARITY.
FT DISULFID 16 39 BY SIMILARITY.

FT DISULFID 41 53 BY SIMILARITY.
FT DISULFID 54 59 BY SIMILARITY.
FT SITE 45 47 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 61 AA; 7024 MW; 34B25A68D48F7EAA CRC64;

Query Match 9.1%; Score 46; DB 1; Length 61;
Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 54 CRNPDGDSGPMWC 56
Dl : : | | | : | : | : | | : |
Db 41 CFTPRGDMGPGYC 53

RESULT 15
YG19_VIBVU
ID YG19_VIBVU STANDARD; PRT; 64 AA.
AC Q8DG30;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical zinc-binding UPF0243 protein VV11619.
GN VV11619.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- COFACTOR: Binds 1 zinc ion (By similarity).
CC -!- SIMILARITY: Belongs to the UPF0243 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF016802; AAC10038.1; ALT_INIT.
DR HMAP; MF_00649; -, 1.
KW Hypothetical protein; Zinc; Metal-binding; Complete proteome.
FT METAL 9 9 ZINC (BY SIMILARITY).
FT METAL 12 12 ZINC (BY SIMILARITY).
FT METAL 28 28 ZINC (BY SIMILARITY).
FT METAL 32 32 ZINC (BY SIMILARITY).
SQ SEQUENCE 64 AA; 7240 MW; AB06E7809FFF59FE CRC64;

Query Match 9.1%; Score 46; DB 1; Length 64;
Best Local Similarity 27.0%; Pred. No. 1.3e+02;
Matches 17; Conservative 3; Mismatches 23; Indels 20; Gaps 3;

QY 18 TWVTGTTCQ-----EWAAQEPHRHSVFIFGTNK-----WAGLEKNCYCRNPDGIN 62
Dl : : | | | | | : | : | : | | : | : |
Db 2 TKIIVKPCQGTNWVGEQSPHR-----PFCSKKQCIDFGWEADENAIIPGAPMDS 56

QY 63 GFV 65
Dl :
Db 57 DGW 59

Search completed: September 29, 2003, 08:47:29
Job time : 24 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 08:44:01 ; Search time 94 seconds
(without alignments)
236.091 Million cell updates/sec

Title: US-10-088-548-8

Perfect score: 508

Sequence: 1 EDCMFGNGKGYRKKATV.....YTMNPKLFDCDIPICASS 86

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 105943

Minimum DB seq length: 0

Maximum DB seq length: 86

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.podent:*
- 12: sp.virus:*
- 13: sp.vertebrate:*
- 14: sp.unclassified:*
- 15: sp.rvirus:*
- 16: sp.bacteriap:*
- 17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	130	25.6	75	6 Q9BGN9	Q9bgn9 bos taurus
2	127.5	25.1	60	4 Q9UKJ7	Q9ukj7 homo sapien
3	123	24.2	53	4 Q9UMI2	Q9umi2 homo sapien
4	56	11.0	67	2 Q9FV6	Q9fov6 azoarcus sp
5	54.5	10.7	65	16 Q9KPE1	Q9kpel vibrio chol
6	54	10.6	25	4 Q9UD88	Q9ud88 homo sapien
7	53	10.4	84	5 Q9W0Y8	Q9w0y8 drosophila
8	51.5	10.1	71	5 Q8MQD3	Q8mqd3 caenorhabdi
9	51	10.0	65	12 Q55745	Q55745 chilo iride
10	50	9.8	48	3 Q8TGR4	Q8tgr4 saccharomyc
11	49.5	9.7	51	10 Q8GRV2	Q8grv2 oryza sativ
12	49.5	9.7	68	16 Q9P8P2	Q9p8p2 xyella fas
13	48.5	9.5	61	15 Q89143	Q89143 visna virus
14	48.5	9.5	85	5 Q9ELM4	Q9elm4 androctonus
15	48.5	9.5	86	15 Q8ADD0	Q8add0 human immun
16	48	9.4	38	5 Q8WP57	Q8wp57 lutzomyia l

17	48	9.4	38	5 Q8WRR7	Q8wrr7 lutzomyia u
18	48	9.4	38	5 Q8WRR5	Q8wrr5 lutzomyia e
19	48	9.4	38	5 Q8WRR9	Q8wrr9 lutzomyia i
20	48	9.4	38	5 Q8WRR2	Q8wrr2 lutzomyia m
21	48	9.4	38	5 Q8WRR4	Q8wrr4 lutzomyia l
22	48	9.4	38	5 Q8WRR3	Q8wrr3 lutzomyia d
23	48	9.4	38	5 Q8WRR8	Q8wrr8 lutzomyia w
24	48	9.4	38	16 Q8XRS9	Q8xrs9 ralstonia s
25	47.5	9.4	28	3 Q9URD3	Q9urd3 aspergillus
26	47.5	9.4	59	2 Q44448	Q44448 agrobacteri
27	47.5	9.4	67	16 Q9KOP3	Q9kop3 neisseria m
28	47.5	9.4	78	15 Q993N1	Q993n1 human immun
29	47	9.3	60	15 Q89146	Q89146 visna virus
30	47	9.3	64	15 Q89145	Q89145 visna virus
31	47	9.3	67	11 Q9ER13	Q9er13 rattus norv
32	47	9.3	77	16 Q8YK17	Q8yki7 anabaena sp
33	47	9.3	86	10 Q9LQ96	Q9lq96 oryza sativ
34	47	9.3	86	11 Q8CJA5	Q8cja5 mus musculu
35	46.5	9.2	51	16 Q8YHK9	Q8yhk9 brucella me
36	46.5	9.2	62	17 Q26228	Q26228 methanobact
37	46.5	9.2	65	15 Q89144	Q89144 visna virus
38	46.5	9.2	73	16 Q9KD21	Q9kd21 bacillus ha
39	46.5	9.2	85	5 Q9GUA7	Q9gua7 mesobuthus
40	46.5	9.2	85	5 Q9SP69	Q9sp69 mesobuthus
41	46.5	9.2	86	4 Q9UFL9	Q9ufl9 homo sapien
42	46.5	9.2	86	15 Q8AD74	Q8ad74 human immun
43	46	9.1	71	12 Q85298	Q85298 orf virus.
44	46	9.1	75	11 Q8R043	Q8r043 mus musculu
45	45.5	9.0	52	12 Q92320	Q92320 human respi

ALIGNMENTS

RESULT 1

Q9BGN9	ID	Q9BGN9	PRELIMINARY;	PRT;	75 AA.
AC	Q9BGN9;				
DT	01-JUN-2001 (TREMBLrel. 17, Created)				
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)				
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)				
DE	Hepatocyte growth factor (Fragment).				
GN	HGF.				
OS	Bos taurus (Bovine).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Bovidae; Bovinae; Bos.				
OX	NCBI_TaxID=9913;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Uterus;				
RA	Murakami S., Fujiwara C., Miyamoto Y., Takeuchi S., Takahashi S.,				
RA	Okuda K.;				
RT	"Expression and action of hepatocyte growth factor in bovine				
RT	endometrial stromal and epithelial cells in vitro.";				
RL	Submitted (PDB-2001) to the EMBL/GenBank/DBJ databases.				
CC	!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.				
DR	EMBL; AB056447; BAB33031.1; -.				
DR	HSSP; P14210; LBHT.				
DR	InterPro; IPR000001; Kringle.				
DR	Pfam; PF00051; Kringle; 2.				
DR	PRINTS; PR00018; KRINGLE.				
DR	ProDom; PD000395; Kringle; 2.				
DR	SMART; SM00130; KR; 2.				
DR	PROSITE; PS00021; KRINGLE_1; 1.				
DR	PROSITE; PS50070; KRINGLE_2; 2.				
KW	Glycoprotein; Kringle.				
FT	NON_TER 1				
FT	NON_TER 75				
FT	NON_TER 75				
SQ	SEQUENCE 75 AA; 8831 MW; 829EEFCC49701B1 CRC64;				
	Query Match 25.6%; Score 130; DB 6; Length 75;				
	Best Local Similarity 57.6%; Pred. No. 1.3e-07;				

```

Matches 19; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 52 NYCNRPDGSDINGPWCYTMNPKLFYCDIPLCA 84
    ||||| 1: ||||| 1: ||||| 1: ||||| 1:
Db 1 NYCNRPDGEGGFWCFISNPEVRYCVDIPLQCS 33

RESULT 2
Q9UKJ7 PRELIMINARY; PRT; 60 AA.
AC Q9UKJ7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21181705; PubMed=11285247;
RA Ogorekova M., Kraft H.G., Enholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringle IV
RT types 6 to 10 domain affect lip(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158657; AAF03676.1; -.
DR HSSP; P00747; IKRN.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00031; Kringle; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 6799 MW; 5719AA26B3E0FFID CRC64;

Query Match 25.1%; Score 127.5; DB 4; Length 60;
Best Local Similarity 55.6%; Pred. No. 2e-07;
Matches 20; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

QY 48 GLENYCRNPDGSDINGPWCYTMNPKLFYCDIPLC 83
    || ||||| 1: ||||| 1: ||||| 1: ||||| 1:
Db 1 GLTNYCRNPDGEIR-PCYTMNPDPSVWFYCNLTQC 35

RESULT 3
Q9UMI2 PRELIMINARY; PRT; 53 AA.
AC Q9UMI2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE PLG protein (fragment).
GN PLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RT human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; K02921; AAA60123.1; -.

```

```

DR HSSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; FALSE_NEG.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Glycoprotein; Kringle.
FT NON_TER 1
FT NON_TER 53
SQ SEQUENCE 53 AA; 5893 MW; 2EB36E715CD40A47 CRC64;

Query Match 24.2%; Score 123; DB 4; Length 53;
Best Local Similarity 54.3%; Pred. No. 5,9e-07;
Matches 19; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 2 QDCMFGNGKGYRGKATVTGTCQENWAAQPHRH 36
    ||| 1: ||| 1: ||| 1: ||| 1: ||| 1:
Db 9 QDCYHGDQSYRGTSSTTTTGKKQCSWSMTPHRH 43

RESULT 4
Q9F0V6 PRELIMINARY; PRT; 67 AA.
AC Q9F0V6;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Electron carrier protein ferredoxin N.
GN FXN.
OS Azoarcus sp. BH72.
OC Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales;
OC Rhodocyclaceae; Azoarcus.
OX NCBI_TaxID=62928;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BH72;
RX MEDLINE=21264388; PubMed=11371540;
RA Egner T., Martin D.E., Sarkar A., Reinhold-Hurek B.;
RT "Role of a Ferredoxin Gene Cotranscribed with the nifHDK Operon in
RT N(2) Fixation and Nitrogenase 'Switch-Off' of Azoarcus sp. Strain
RT BH72.";
RL J. Bacteriol. 183:3752-3760(2001).
DR EMBL; AF200742; AAG35589.1; -.
DR HSSP; P00193; IDUR.
DR InterPro; IPR001450; 4Fe4S_ferredoxin.
DR Pfam; PF00037; fer4; 1.
DR PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.
KW 4Fe-4S; Iron; Iron-sulfur.
SQ SEQUENCE 67 AA; 6997 MW; 8E5A0E4A6FEE157B CRC64;

Query Match 11.0%; Score 56; DB 2; Length 67;
Best Local Similarity 33.3%; Pred. No. 44;
Matches 12; Conservative 6; Mismatches 16; Indels 2; Gaps 1;

QY 49 LEKNYCRNPDGSDINGPWCYTMNPKLFYCDIPLCA 84
    : 1: ||| 1: ||| 1: ||| 1: ||| 1:
Db 34 INKDYCTCEGDFDEPQCLAVCPAG--DACIVPLAA 67

RESULT 5
Q9KPE1 PRELIMINARY; PRT; 65 AA.
AC Q9KPE1;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hypothetical protein VC2429.
GN VC2429.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;

```



```

RN RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10953301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Hart D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uytterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004312; AAF95572.1; -.
DR TIGR; VC2429; -.
DR InterPro; IPR005584; DUF329.
DR Pfam; PF03884; DUF329; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 65 AA; 7397 MW; 2ED867C1D1D8AD90 CRC64;

Query Match 10.7%; Score 54.5; DB 16; Length 65;
Best Local Similarity 31.7%; Pred. No. 64;
Matches 20; Conservative 2; Mismatches 24; Indels 17; Gaps 3;

QY 15 KKATTVTGTPC--QEWAAQPHRHSTFIPGINK-----WAGLEKYNCRNPDGGIN 62
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3 KLUITVKPCFGTGVWGEQSPHR-----PFCSKOCQIMDFGEWADEKAIPGAPMDS 57

QY 63 GPW 65
Db 58 DGW 60

RESULT 6
Q9UD88 PRELIMINARY; PRT; 25 AA.
AC Q9UD88;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE APO(A) KRINGLE 4-37, APO(A) KRINGLE type 5 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94060120; PubMed=7848387;
RA Pfaffinger D., McLean J., Scaru A.M.;
RT "Amplification of human APO(a) kringle 4-37 from blood lymphocyte
RT DNA.";
RL Biochim. Biophys. Acta 1225:107-109(1993).
DR HSP; P00747; IKNR.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
KW Glycoprotein; Kringle.
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA; 2928 MW; BEEDD4C62FA480A8 CRC64;

Query Match 10.6%; Score 54; DB 4; Length 25;
Best Local Similarity 33.3%; Pred. No. 25;
Matches 7; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 66 CYTWNPKRKYDIDIPLCASS 86
Db ||||| :|||: ||:
1 CFTMDPSIRWEYCNLRCSDT 21

RESULT 7
Q9W0Y8 PRELIMINARY; PRT; 84 AA.
ID Q9W0Y8

```

```

AC Q9W0Y8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG9071 protein.
GN NACP608 OR CG9071.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkelley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazef R.G., Champé M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Bailley R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berzhan B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslér C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Welstock G.M., Weissbach J.,
RA Williams S.M., Woodge T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2165-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Amanatides P.G., Brandon R.C., Rogers Y., An H., Baldwin D.,
RA Banzon J., Beeson K.Y., Busam D.A., Carlson J.W., Center A.,
RA Champe M., Davenport L.B., Dietz S.M., Dodson K., Dorsett V.,
RA Doup L.E., Doyle C., Dresnek D., Farfan D., Ferreira S., Frise E.,
RA Galle R.F., Garg N.S., George R.A., Gonzalez M., Houck J.,
RA Hoskins R.A., Hostin D., Howland T.J., Ibegwam C., Jalali M.,
RA Kruse D., Li P., Mattei B., Moshrefi A., McIntosh T.C., Moy M.,
RA Murphy B., Nelson C., Nelson K.A., Nunoo J., Pacleb J., Paragas V.,
RA Park S., Patel S., Pfeiffer B., Phouanavong S., Pittman G.S.,
RA Puri V., Richards S., Scheeler F., Stapleton M., Strong R.,
RA Svirkas R., Tector C., Tyler D., Williams S.M., Zaveri J.S.,
RA Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,

```

```

RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celniker S.E.,
RA Clamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
RA Harris N.L., Krommiller B., Marshall B., Millburn G.H., Richter J.,
RA Russo S., Searle S.M.J., Smith E., Shu S., Smtunak F.,
RA Whitfield E.J., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003465; AAF47291.3; -.
DR FlyBase; FBgn0002920; NACP60B.
SQ SEQUENCE 84 AA; 9648 MW; 8668C9865252B777 CRC64;

Query Match 10.4%; Score 53; DB 5; Length 84;
Best Local Similarity 25.4%; Pred. No. 1.3e+02;
Matches 16; Conservative 9; Mismatches 18; Indels 20; Gaps 3;

QY 13 RKKKATVTGPPCEW-----AAQEPHRSHTFPGT-----NKGWGLEKNYC 54
|||: ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 22 RGRRLSTRGT--SSWGGSTASTRALPPWRTATCCPANGSPFRSTCTSRWRRLTFSFT 79
|||: ||| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 55 RNP 57
| | | | |
Db 80 RRP 82
| | | | |

RESULT 8
Q8MQD3 PRELIMINARY; PRT; 71 AA.
ID Q8MQD3;
AC Q8MQD3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein C18A3.10.
GN C18A3.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peleoderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Hallsworth K.;
RT "The sequence of C. elegans cosmid C18A3.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; U28944; AAM61117.1; -.
DR WormPep; C18A3.10; CE31306.
KW Hypothetical protein.
SQ SEQUENCE 71 AA; 8102 MW; 2C3669D844A9A3E3 CRC64;

Query Match 10.1%; Score 51.5; DB 5; Length 71;
Best Local Similarity 26.9%; Pred. No. 1.6e+02;
Matches 14; Conservative 6; Mismatches 19; Indels 13; Gaps 2;

QY 16 KATVTGTGTCQWAAQEPHRSHTFPGTNGWGLEKNYCRNPDGNGPWCY 67
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 5 KATPTVSTIESGTDPHEHYRY-----REKNIC-----GISGPICF 43
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
O55745

```

```

ID O55745 PRELIMINARY; PRT; 65 AA.
AC O55745;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Hypothetical 7.5 kDa protein.
OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).
OS Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=10488;
RN [1]
RP SEQUENCE FROM N.A.
RA Bahr U., Tidona C.A., Darai G.;
RL Virus Genes 0:0-0(1997).
DR EMBL; AF303741; AAB94456.1; -.
KW Hypothetical protein.
SQ SEQUENCE 65 AA; 7526 MW; 1DCD213269477488 CRC64;

Query Match 10.0%; Score 51; DB 12; Length 65;
Best Local Similarity 29.1%; Pred. No. 1.6e+02;
Matches 16; Conservative 4; Mismatches 19; Indels 16; Gaps 2;

QY 39 FIPGTNKGWGLEKNYCRNPDGNGPWCY-----TMNPKLFYCDIPLC 83
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 14 FYTLTNVFAFPDGMCLNLDGVSQYGVFRRVESPTTTKN-----YCDYYC 62
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 10
Q8TGR4 PRELIMINARY; PRT; 48 AA.
ID Q8TGR4;
AC Q8TGR4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 5.2 kDa protein.
GN Y5R088W-A.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21624570; PubMed=11753363;
RA Kumar A., Harrison P.M., Cheung K.H., Ian N., Echols N., Bertone P.,
RA Miller P., Gerstein M.B., Snyder M.;
RT "An integrated approach for finding overlooked genes in yeast.";
RL Nat. Biotechnol. 20:58-63(2002).
DR EMBL; AF479921; AAL79234.1; -.
KW Hypothetical protein.
SQ SEQUENCE 48 AA; 5208 MW; F7C3E1293C6B0202 CRC64;

Query Match 9.8%; Score 50; DB 3; Length 48;
Best Local Similarity 35.7%; Pred. No. 1.5e+02;
Matches 10; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

QY 53 YCRNPDG--DINGPWCYTMNPKLFYDC 78
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 20 YCHSRDGRKPGHGWTFGAVGPRQTGECC 47
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
Q8GRV2 PRELIMINARY; PRT; 51 AA.
ID Q8GRV2;
AC Q8GRV2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE P0438G07.15 protein (B1126E07.15 protein).
GN P0438G07.15 OR B1126E07.15
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartidae; Oryzae; Oryza.
OX NCBI_TaxID=39947;

```

RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, PAC
 clone:PO438G07."
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
 clone:B1126F07."
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP004398; BAC22372.1; -
 DR EMBL: AP005674; BAC22483.1; -
 SQ SEQUENCE 51 AA; 5702 MW; B4EC539E549BF6C CRC64;

 Query Match 9.7%; Score 49.5; DB 10; Length 51;
 Best Local Similarity 29.4%; Pred. No. 1.8e+02;
 Matches 15; Conservative 8; Mismatches 27; Indels 1; Gaps 1;

 QY 5 MFGNGKGVGCKATVGTGTCQWAAQEPHRSSTFIPGTNKWAGLEKNYCR 55
 DB 1 MGEGRRRRRGEGEETIGGE-EEERGSEGRGDEIDPMLSKYPGEESCR 50

 RESULT 12
 Q9BPB2 PRELIMINARY; PRT; 68 AA.
 ID Q9BPB2
 AC Q9BPB2
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein Xf2098.
 GN Xf2098.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9a5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Canargo A., Camargo L.E.A., Carraro D.M., Carrier H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.B., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramse E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva P.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.F., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa."
 RL Nature 406:151-159(2000).
 DR EMBL: AE004025; AAF84897.1; -.

KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 68 AA; 7349 MW; 87902D67E5522F41 CRC64;

 Query Match 9.7%; Score 49.5; DB 16; Length 68;
 Best Local Similarity 31.6%; Pred. No. 2.5e+02;
 Matches 12; Conservative 3; Mismatches 6; Indels 17; Gaps 2;

 QY 65 WCYTNPR-----KLFDY-----CDIPLCAS 85
 DB 5 WCYQSTRAAISLQELALRLQYRAAFMSCALPLCAS 42

 RESULT 13
 Q89143 PRELIMINARY; PRT; 61 AA.
 ID Q89143
 AC Q89143
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Envelope polyprotein (Fragment).
 GN ENV.
 OS Vigna virus.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11741;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=676;
 RX MEDLINE=95302052; PubMed=7782784;
 RA Leroux C., Vuillermoz S., Mornex J.F., Greenland T.;
 RT "Genomic heterogeneity in the pol region of ovine lentiviruses
 obtained from bronchoalveolar cells of infected sheep from France."
 RL J. Gen. Virol. 76:1533-1537(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=676;
 RX MEDLINE=97372948; PubMed=9229003;
 RA Leroux C., Chastang J., Greenland T., Mornex J.F.;
 RT "Genomic heterogeneity of small ruminant lentiviruses: existence of
 heterogeneous populations in sheep and of the same lentiviral
 genotypes in sheep and goats."
 RL Arch. Virol. 142:1125-1137(1997).
 DR EMBL: U35858; AAA78287.1; -.
 KW Polyprotein.
 FT NON_TER 1
 FT NON_TER 61
 SQ SEQUENCE 61 AA; 7240 MW; C88726FF6FB078B0 CRC64;

 Query Match 9.5%; Score 48.5; DB 15; Length 61;
 Best Local Similarity 26.5%; Pred. No. 2.9e+02;
 Matches 13; Conservative 7; Mismatches 18; Indels 11; Gaps 2;

 QY 26 QEWAA-----QPHRHSSTFIPGTNKWAGLEKNY-CRNPDGDDING 63
 DB 13 QQWTC SARVKWKRIQELKDSLYIAGDFWGVKAKYSCERNIGRLDG 61

 RESULT 14
 Q9BLM4 PRELIMINARY; PRT; 85 AA.
 ID Q9BLM4
 AC Q9BLM4
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative toxin precursor.
 OS Androctonus australis (Sahara scorpion).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
 OC Buthoidea; Buthidae; Androctonus.
 OX NCBI_TaxID=6858;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21211344; PubMed=11311249;
 RA Ceard B., Martin-Eauclaire M.F., Bougis P.E.;
 RT "Evidence for a position-specific deletion as an evolutionary link

```

RT between long- and short-chain scorpion toxins." ;
RL FEBS Lett. 494:246-248(2001).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR EMBL: AJ308440; CAC37321.1; -.
DR HSP; P01484; IAHO.
DR InterPro: IPR003614; Knot1.
DR InterPro: IPR001219; Neurotoxin.
DR InterPro: IPR002061; Scorpion_toxinL.
DR Pfam: PF00537; toxin_3; 1.
DR PRINTS: PR00284; TOXIN.
DR ProDom: PD000908; Scorpion_toxinL; 1.
DR SMART: SM00505; Knot1; 1.
KW Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 85 PUTATIVE TOXIN.
SQ SEQUENCE 85 AA; 9523 MW; 4059A69D80E4F090 CRC64;

Query Match 9.5%; Score 48.5; DB 5; Length 85;
Best Local Similarity 31.9%; Pred. No. 4.3e+02;
Matches 15; Conservative 6; Mismatches 15; Indels 11; Gaps 3;

CY 35 RHSTFIFGTN-----KWAGLEKNYCR--NPGDINGPCYTMNPR 72
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 29 KNCTEFCGRNAYCNDECKKKGAEISGYCOWASPYG--NACYCYKLPDR 73

RESULT 15
Q8ADD0 Q8ADD0 PRELIMINARY; PRT; 86 AA.
AC Q8ADD0;
RX MEDLINE=22375625; PubMed=12487816;
RA Harris M.E., Serwadda D., Sewankambo N., Wabwire F., Kim B.,
RA Kigozi G., Kiwanuka N., Phillips J.B., Meehen M., Lutalo T.,
RA Laue J.R., Merling R., Gray R., Wawer M., Birx D.L., Robb M.L.,
RA McCutchan F.E.;
RT "Among 46 Near Full Length HIV Type 1 Genome Sequences from Rakai
RT District, Uganda, Subtype D and AD Recombinants Predominate.";
RL AIDS Res. Hum. Retroviruses 18:1281-1290(2002).
DR EMBL: AF484515; AAN73775.1; -.
SQ SEQUENCE 86 AA; 9726 MW; DBCABIAAFDBD13B8 CRC64;

Query Match 9.5%; Score 48.5; DB 15; Length 86;
Best Local Similarity 27.7%; Pred. No. 4.3e+02;
Matches 18; Conservative 4; Mismatches 26; Indels 17; Gaps 2;

QY 2 QDCMFNGKGG--YRGKATVTGTGCPQEWAAQEPHRSHTFIPGNKWAGLEKNYCRNPDG 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 35 QNCFTKGLGISYGRKKRRKRTKPGDPANQD-----IPKPSQPRG 79

QY 60 DINGP 64
   |||
DB 80 DSTGP 84

```

Search completed: September 29, 2003, 08:49:13
Job time : 98 secs